

University of Arkansas, Fayetteville ScholarWorks@UARK

Theses and Dissertations

12-2013

The Effects of Social Contextual Factors on Rater Motivation and Performance Ratings

Emilija Djurdjevic

University of Arkansas, Fayetteville

Follow this and additional works at: <http://scholarworks.uark.edu/etd>

 Part of the [Quantitative, Qualitative, Comparative, and Historical Methodologies Commons](#)

Recommended Citation

Djurdjevic, Emilija, "The Effects of Social Contextual Factors on Rater Motivation and Performance Ratings" (2013). *Theses and Dissertations*. 943.

<http://scholarworks.uark.edu/etd/943>

This Dissertation is brought to you for free and open access by ScholarWorks@UARK. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of ScholarWorks@UARK. For more information, please contact scholar@uark.edu, ccmiddle@uark.edu.

The Effects of Social Contextual Factors on Rater Motivation and Performance Ratings

The Effects of Social Contextual Factors on Rater Motivation and Performance Ratings

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor in Philosophy in Business Administration

By

Emilija Djurdjevic
University of South Florida
Bachelor of Arts in Psychology, 2008

December 2013
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Dr. Christopher C. Rosen
Dissertation Chair

Dr. Nina Gupta
Committee Member

Dr. John E. Delery
Committee Member

ABSTRACT

Performance appraisal research (e.g., Longenecker, Sims, & Gioia, 1987; Wang, Wong, & Kwong, 2010) suggests that inaccuracies in performance ratings are largely a result of intentional rating distortion, or rater motivation. Researchers have noted that raters sometimes have multiple and conflicting goals when providing performance ratings (e.g., Murphy, Cleveland, Kinney, Skattebo, Newman, & Sin, 2003). According to rater goal theory (e.g., Cleveland & Murphy, 1992), rater goals are influenced by the performance appraisal context, mediating the relationship between the context of performance appraisal and performance ratings. Studies suggest that the context affects performance ratings, but conceptualizations of the performance appraisal context (e.g., rating purpose) have been narrow, neglecting the social aspects inherent in performance appraisal (Kozlowski, Chao, & Morrison, 1998; Levy & Williams, 2004). As such, we lack a clear understanding of what the key social contextual factors are and how they affect rater motivation and performance ratings.

To design effective interventions to improve rating accuracy, we first need to understand the strength and direction of these influences (Spence & Keeping, 2010). Therefore, the present study considers how key social contextual factors (i.e., the rater's supervisor and subordinate group) independently and simultaneously affect rater motivation and performance ratings. Expectancy theory (Porter & Lawler, 1968; Vroom, 1964) is used to make predictions about the effects of the rater's supervisor and group on rater motivation and performance ratings. Expectancy theory is then integrated with social impact theory (Latané, 1981) to develop hypotheses regarding their interactive effects. Furthermore, both rating inflation and rating deflation, which has received limited research attention, are explicitly addressed. The results of the present study provide some evidence that the key social contextual factors interact to

influence rater motivation and performance ratings. Additionally, the results provide empirical support for different effects of information encouraging high versus low performance ratings. The findings are discussed in detail, along with theoretical and practical implications, and directions for future research.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincerest gratitude to my dissertation chair and mentor, Chris Rosen, for his dedication, patience, advice, and feedback over these last five years. He has been invaluable to my development as an academic. I would also like to thank Nina Gupta, my committee member, for her continued commitment and generosity. Additionally, I want to thank John Delery for serving on my committee and for his help and insightful feedback. Finally, I would like to thank Russ Johnson, who introduced me to research and academia.

Graduate school would not have been the same without the friends I made along the way. I would like to acknowledge my cohort, Ru-Shiun Liou and Amy Guerber, who have always kept me motivated. I am also grateful for two amazing friends and colleagues, Samantha Conroy and Shannon Rawski, with whom I shared the graduate school experience. I would also like to extend my thanks to the individuals who have contributed to this study, including Daanish Pestonjee, Sarah Holtzen, the expert raters, and others who helped with the data collection process and study.

Finally, I would like to acknowledge my family. I am thankful for the encouragement my parents, Branislav and Zlata Djurdjevic, have provided throughout the dissertation process. Lastly, I want to thank my sister, Mijana Djurdjevic, who has always been there for me. I would not be where I am without the support of these individuals.

TABLE OF CONTENTS

I. STATEMENT OF THE PROBLEM	1
II. REVIEW OF THE LITERATURE	9
Performance Appraisal (PA)	9
The Importance of Rating Accuracy	9
Improving Rating Accuracy	10
The Role of Rater Motivation	12
The Performance Appraisal Context	15
Key Contextual Factors in PA	19
Theoretical Frameworks	25
Expectancy Theory	25
Integrating Expectancy Theory with the Social Context of PA	28
Social Impact Theory	32
Integrating Social Impact Theory with the Social Context of PA	35
Social Impact Theory, Expectancy Theory, and the PA Context	37
Present Study	39
Overview	39
Forces in the PA Context	39
Hypothesis Development and Hypotheses	42
III. METHODOLOGY	60
Subjects	60
Task and Materials	60
Development of Rating Scales	61

Development and Taping of Performance Scenarios	62
Validation of True Score Estimates	62
Dependent Variables	63
Procedures	64
Experimental Manipulations	66
Manipulation Checks	66
IV. RESULTS	75
Evaluation of Interrater Agreement	75
Tests of Hypotheses	78
3 x 3 Factorial ANOVA	78
Inflation versus Deflation Information	80
Main Effects of Supervisor	81
Main Effects of Subordinate Group	82
Interaction Effects of Supervisor and Subordinate Group	85
Follow-Up Analyses	101
Effects of Manipulations on Desire to Inflate or Deflate Ratings	101
Effects of Manipulations on $P \rightarrow O$ Expectancies	105
Effects of Manipulations on Valences	106
Interaction Effects of Manipulations	108
Interaction Effects of Manipulations on Valences	108
Interaction Effects of Manipulations on $P \rightarrow O$ Expectancies	110
V. DISCUSSION	113
Summary of Supervisor Main Effects	117

Summary of Subordinate Group Main Effects	118
Summary of Supervisor and Subordinate Group Interaction Effects	120
Summary of Effect Magnitudes	122
Theoretical and Practical Implications	122
Limitations and Directions for Future Research	132
Summary	139
VI. REFERENCES	141
VI. APPENDICIES	152
Conceptual Model	152
Description of Participant's Role as Bank Manager	153
Scripts for Practice Performance Scenarios	154
Scripts for Performance Rating Scenarios	157
Performance Rating Scale	165
Behaviorally-Anchored Rating Scales	166
Graphic Rating Scale	169
Formulas for Calculating Rating Distortion	170
Overview of Study for Participants	171
Experimental Manipulations	172
Formulas for Comparing Effect Magnitudes for ANOVA Designs	174
Research Compliance Protocol Letters	175

CHAPTER 1

STATEMENT OF THE PROBLEM

The majority of organizations in the United States use performance appraisal (PA), most commonly supervisory performance ratings (Murphy, 2008), to evaluate employee performance. These performance evaluations are used to make administrative decisions, motivate and develop employees, determine training needs, and evaluate recruitment and selection processes (e.g., Campbell, Dunnette, Lawler, & Weick, 1970; Cleveland, Murphy, & Williams, 1989; Kane & Lawler, 1979; Levine, 1986). Despite the popularity and importance of performance appraisal, the relationship between actual performance and subjective ratings of job performance is relatively weak (Murphy, 2008; Viswesvaran, Ones, & Schmidt, 1996). This is problematic, as it suggests that performance evaluations are contaminated with non-performance factors (e.g., impression management: Wayne & Ferris, 1990) and therefore capture less of the relevant criterion space (Austin & Villanova, 1992). As performance ratings do not yield the intended information, the effectiveness of the PA system as well as interrelated organizational systems, such as the pay system, is diminished (Lawler, 1990).

These concerns have led researchers to focus on improving performance appraisal accuracy. Early research focused on reducing rating errors by improving psychometric properties of rating instruments (e.g., Kane & Bernardin, 1982; Latham & Wexley, 1977). This was followed by addressing specific rating errors (e.g., halo) and providing raters with training to reduce these errors (e.g., Athey & McIntyre, 1987; McIntyre, Smith, & Hassett, 1984). Subsequent performance appraisal research was largely influenced by Landy and Farr's (1980) seminal paper on the cognitive processes underlying performance appraisal. This research

centered on how raters acquire, encode, store, and retrieve information about ratees. Together, these approaches to improving performance appraisal accuracy have contributed to our understanding of the rating process. They have failed, however, to lead to significant progress in the PA literature and provide practical recommendations for practitioners (Banks & Murphy, 1985; Ilgen, Barnes-Farrell, & McKellin, 1993; Longenecker, Sims, & Gioia, 1987).

The lack of notable advancement in improving PA accuracy (i.e., strengthening the relationship between job performance and performance ratings) is largely due to the underlying assumption that raters are willing or motivated to provide accurate performance ratings. Scholars (e.g., Mohrman & Lawler, 1983) called attention to rater motivation in performance appraisal thirty years ago. The field of performance appraisal, however, has only slowly shifted to focus on the effects of rater motivation on performance appraisal accuracy. Empirical studies (e.g., Wang, Wong & Kwong, 2010) and interviews with managers (Longenecker et al., 1987) provide evidence that raters are more interested in accomplishing personal goals than providing accurate performance ratings. There is also evidence that supervisors pursue multiple goals that affect employees' performance ratings (e.g., Kozlowski, Chao, & Morrison, 1998; Wong & Kwong, 2007).

It has been suggested that rater goals and rating accuracy are affected by the performance appraisal context (e.g., Cleveland & Murphy, 1992; Harris, 1994). The importance of the performance appraisal context has been acknowledged (DeCotiis & Petit, 1978; Landy & Farr, 1980; Murphy & Cleveland, 1991) and theorized to affect raters' motivation to provide accurate ratings (e.g., Harris, 1994). The performance appraisal context represents situational, or organizational and performance appraisal factors, other than a ratee's performance, which influence performance ratings (Judge & Ferris, 1993; Murphy, 2008). For example, rating

purpose has been shown to have an effect on performance ratings (e.g., Greguras, Robie, Schleicher, & Maynard, 2003). The effects of the social context (i.e., the social intricacies, or other individuals, in the PA context) on rater motivation and rating accuracy have received little empirical research, however (Levy & Williams, 2004). The limited research that has been conducted in the area suggests that social factors in the PA context, such as the extent to which others inflate performance ratings or an employee's influence tactics, affect performance ratings (Judge & Ferris, 1993; Spence & Keeping, 2010).

The dearth of empirical research in this area leaves us without a clear understanding of what the key social contextual factors are and how they affect rater motivation and performance ratings. Spence and Keeping (2013) have noted that we do not fully understand the factors raters consider and “what determines their relative strength/intensity” (p. 2). In order to design effective interventions to improve performance appraisal accuracy, we first need to identify key social factors affecting performance ratings and develop theory that identifies how this occurs. Then, we need to test this theory to empirically establish the strength and direction of the influences (Spence & Keeping, 2010). This will inform our understanding of how the complex interpersonal relationships inherent in the performance appraisal context, which have been largely neglect in the literature, influence rating accuracy (Levy & Williams, 2004).

Given the social nature of performance appraisal, those with whom the rater has the most frequent contact with should represent key social contextual influences in the rating process. In the workplace, individuals interact with their supervisor and subordinate group, and these individuals motivate employees to engage in any number of behaviors at work (Katz & Kahn, 1978). In fact, research (Kozlowski et al., 1998) suggests that the organizational members the rater reports to, and those who report to the rater, are factors the rater considers when rendering

performance ratings. This suggests that the rater's supervisor and subordinate group should be salient sources of influence in the performance appraisal context when they provide information regarding appropriate and expected behaviors (Reno, Cialdini, & Kallgren, 1993). Therefore, the focus of the present study is on the rater's supervisor and subordinate group as they are the most proximal individuals to the rater and represent individuals the rater interacts with most frequently. The study considers the effects of these social contextual factors in isolation, as well as their interactive effects, on rater motivation and ratings. This represents an important contribution to the PA literature as it explicates and tests the causal and mediating mechanisms between the key social contextual factors in performance appraisal and performance ratings. A conceptual model for this study is illustrated in Figure 1 in Appendix A¹.

Performance evaluations determine the allocation of critical and scarce resources (e.g., pay raises and promotions). Research suggests that when scarce and critical resources are being allocated, decision makers encounter social influences intended to affect their decisions (Pfeffer, Salancik, & Leblebici, 1976; Zalesny, 1990). It is therefore likely that raters will encounter information from both the supervisor and group when conducting performance appraisals. Social influences can be conceptualized as instances where the rater encounters behaviors, instructions, or other information from an individual or a group that are intended to change his or her behavior to the influencer's desired end (Sussman & Vecchio, 1982). Given the implications and subjective nature of performance evaluations, social influences are an inherent part of the performance appraisal context. It is therefore important to gain a clear understanding of the main and interactive effects of key social contextual influences (i.e., the rater's supervisor and

¹ While ratee performance and rater ability should also influence performance ratings, the focus of this study is on the social context of performance appraisal. Therefore, the effects of the rater's supervisor and group, who represent key social contextual influences, are examined.

subordinate group) on performance ratings. Examining how information from these sources simultaneously affects rater motivation and performance ratings is a necessary first step in understanding how to improve the accuracy of performance ratings. As such, the present study focuses on developing and testing theory that explains the main and interactive effects of the rater's supervisor and subordinate group on rater motivation and rating accuracy.

To hypothesize about the effects of information from the rater's supervisor and subordinate group on rater motivation and rating accuracy, I use expectancy theory. Mohrman and Lawler (1983) proposed applying expectancy theory (Porter & Lawler, 1968; Vroom, 1964) to understand rater motivation, and researchers (e.g., DeCotiis & Petit, 1978; Mohrman & Lawler, 1983; Murphy & Cleveland, 1995; Napier & Latham, 1986) have drawn on expectancy theory to develop models of rater motivation and PA accuracy. Expectancy theory allows integration of multiple, and possibly conflicting courses of action, and the various outcomes associated with the actions in a meaningful way (Porter & Lawler, 1968). The theory also permits a more fine-grained understanding of how various motivational components are affected by contextual influences and how each component affects rater motivation to provide accurate performance ratings. This diagnostic approach is useful for practical recommendations (Sanchez, Truxillo, & Bauer, 2000), as specific motivational components that affect outcomes can be isolated. Finally, expectancy theory has been shown to be a valid theory for predicting work behavior (Van Eerde & Thierry, 1996), such as performing routine tasks (e.g., Sheridan, Slocum, & Min, 1975).

According to expectancy theory, individuals consider future outcomes associated with each potential behavior when choosing among alternative behavioral options (Lawler, 1973). For example, if providing inflated performance ratings results in overall more valuable expected

outcomes for the rater (e.g., promotions and positive interpersonal relationships) than providing accurate performance ratings, the rater will provide inflated performance ratings. Alternatively, if providing deflated performance ratings results in more valuable outcomes overall than providing accurate performance ratings, the rater will provide deflated performance ratings. Ultimately, the behavior that is most likely to maximize self-interest will have the highest motivational force and will be the behavior an individual chooses to perform (Pinder, 1998). Expectancy theory does not inform our understanding of the valences associated with the various outcomes, and therefore the relative importance of information from the rater's supervisor and subordinate group. To understand the relative importance of information from these different sources on rating accuracy, and what happens when it is encountered by raters simultaneously, I draw on social impact theory (e.g., Latané, 1981). Social impact theory provides a parsimonious theoretical framework that explains the magnitude of effects from sources relatively higher and lower in power than the target (Latané & Wolf, 1981), such as the rater's supervisor and group, respectively.

Based on social impact theory and expectancy theory, and drawing from the social influence literature, I consider the effects of information from the rater's supervisor and subordinate group on rater motivation to provide accurate versus distorted performance ratings and develop predictions regarding ratings. To this end, the study makes four primary contributions to the literature. First, it provides a test of the main and interactive effects of the rater's supervisor and subordinate group, which aids in informing our understanding of how key social contextual factors affect rater motivation and rating accuracy in isolation as well as simultaneously. Researchers (e.g., Farr & Levy, 2007; Spence & Keeping, 2010) have noted that we need to understand which social contextual factors make the performance rating process less

effective before we can design interventions to improve rating accuracy. For example, the present study will illuminate whether ratings are less accurate when only the supervisor provides information regarding performance ratings than when the supervisor and group provide inconsistent information. Second, the current study extends theory by integrating expectancy theory with social impact theory to generate predictions about the interactive effects of information from the rater's supervisor and subordinate group on performance ratings. This explicates the underlying mechanisms connecting the social context of performance appraisal to performance ratings, through rater motivation.

Third, the study builds on previous performance appraisal research by examining the effects of information encouraging high and low performance evaluations on performance ratings. This enriches our understanding of how raters respond to different rating information and how it affects rating accuracy, explicitly focusing on rating deflation, which has received limited attention in the literature. The present study also builds on and extends the limited empirical research (e.g., Mero, Guidice, & Brownlee, 2007; Spence and Keeping, 2010) that has examined social contextual influences on performance ratings by providing a more realistic simulation of the performance appraisal context. For example, raters in Mero and colleagues' (2007) study were not aware of the rating preferences of the audiences. To address this limitation, the supervisor's and group's rating preferences are made more explicit in the present study. Additionally, to ensure realism, raters in the present study observe, evaluate, and rate employee performance instead of simply receiving objective performance information, as in Spence and Keeping's (2010) investigation.

Finally, the present study contributes to the social influence literature. A fine grained examination of the effects of information from the group and supervisor on motivation and

behavior is provided. Specifically, the effects of information from both sources on different components of motivation can be isolated. The interactive effects of the group and supervisor are also examined, informing our understanding of how these two social influences simultaneously affect motivation and behavior.

CHAPTER 2

REVIEW OF THE LITERATURE

Performance Appraisal

The Importance of Rating Accuracy

Organizational effectiveness is contingent on individual performance (Gong, Law, Chang, & Xin, 2009). Thus, most organizations in the United States have some form of performance appraisal system in place to evaluate employees. The most common form of performance appraisal is supervisory performance ratings (Murphy, 2008). Supervisory performance ratings are evaluations of the employee's performance over some period of time provided by the individual the employee reports to, usually an immediate supervisor. Supervisory performance evaluations are used for a variety of organizational purposes. For example, important personnel decisions (e.g., promotions and pay raises) are based on supervisory performance evaluations (Cleveland & Murphy, 1995; Heneman & Werner, 2005). Performance appraisals are also used developmentally (e.g., Boswell & Boudreau, 2002) to determine employees' training needs (Levine, 1986) and motivate and develop them (Cleveland et al., 1989). Moreover, organizations use performance ratings as a criterion against which recruitment and selection processes are evaluated (Cleveland et al., 1989; Kane & Lawler, 1979). Increasingly, performance evaluations are also used in litigation where the organization must demonstrate that decisions (e.g., discharges) are based on employee performance rather than non-performance factors (e.g., gender, race, or age: Martin, Bartol, & Kehoe, 2000; Werner & Bolino, 1997).

Given that performance evaluations are used for a variety of important purposes, it is critical to ensure that performance ratings are accurate. Inaccurate performance ratings lead to diminished effectiveness of the performance appraisal system and negate its intended purposes. Issues with performance appraisal accuracy also affect inter-related organizational systems, such as the pay system (Lawler, 1990). For example, if performance ratings are tied to pay raises and ratings are based on non-performance factors, then pay raises are not being used to reward performance but rather other behavior(s). These may include behaviors such as impression management or influence tactics, which have been shown to affect supervisory performance ratings (e.g., Dulebohn, Shore, Kunze, & Dookeran, 2005; Wayne & Kacmar, 1991). Rewarding these and other non-performance factors weakens the connection between performance and pay, and diminishes the effectiveness of merit pay policies (e.g., Heneman & Werner, 2005). Additionally, validation research that is based on performance ratings as the criterion may not provide organizations with valid results if performance ratings are not accurate (Osterman, 2007). Finally, when performance ratings do not accurately reflect employee performance, employees may view the organization as unfair, increasing the likelihood of lawsuits filed against the organization (Dunford & Devine, 1998), among other things.

Improving Rating Accuracy

Scholars have spent decades working on improving performance appraisal quality (Murphy, 2008). Early research focused on reducing rating errors by addressing psychometric properties of rating instruments (e.g., Kane & Bernardin, 1982; Latham & Wexley, 1977). The work in this area led to some improvements in PA accuracy. For example, studies focusing on the number of response categories (e.g., Bernardin, La-Shells, Smith, & Alvares, 1976; Lissitz & Green, 1975) have found that seven response categories, plus or minus two, are most appropriate

to ensure reliability (Landy & Farr, 1980). Nonetheless, scholars (e.g., Cleveland & Murphy, 1992) agreed that psychometric issues were not the largest impediment to accurate performance evaluations. In part, this is because little progress was made in creating an effective alternative to the graphic rating scale (Landy & Farr, 1980). Additionally, results of several meta-analyses (e.g., Bommer, Johnson, Rich, Podsakoff, & MacKenzie, 1995; Heneman, 1986) show that rating format and rating scales have small moderating effects on the relationship between subjective supervisor ratings of job performance and organizational record-based measures. Specifically, “modifications to rating scale formats had only small effects on the reliability, validity, and utility of performance ratings” (Murphy, 2008: p. 150). These results suggest that psychometrically sound rating instruments are not sufficient for ensuring accurate ratings.

Next, scholars worked on improving performance appraisal quality by focusing on specific rating errors (e.g., halo) and providing raters with training (e.g., Athey & McIntyre, 1987; McIntyre et al., 1984) to reduce these errors. Research (e.g., Noonan & Sulsky, 2001; Sulsky & Day, 1994) has shown that frame-of-reference training (Bernardin & Buckley, 1981), for example, is effective in improving the accuracy of performance ratings. Meta-analytic results (Woehr & Huffcutt, 1994), however, show that overall, training raters is only moderately effective in improving the accuracy of performance ratings. Drawing on findings in the literature, Murphy (2008) noted that in part, this is due to the “limited relationship between rater errors and rating accuracy” (p. 150).

Subsequent performance appraisal research was largely influenced by Landy and Farr’s (1980) seminal paper on the cognitive processes underlying performance appraisal. This approach centered on information processing and the research stemming from it enhanced our understanding of how raters acquire, encode, store, and retrieve information about ratees.

Scholars (e.g., Murphy & Cleveland, 1991), however, have noted that while informative, this line of research neglects non-cognitive influences on performance appraisal, such as social factors (Murphy, Cleveland, Kinney, Skattebo, Newman, & Sin, 2003). Additionally, Murphy and Cleveland (1995) have pointed out several problems with the cognitive approach. First, the errors that are being studied may not occur very commonly or frequently and focus on ability factors. Second, the research is not sufficiently focused on application and therefore issues relevant to PA in organizations are not examined. Third, the laboratory studies conducted in this area focus more on judgments (i.e., private evaluations) rather than ratings (i.e., communicated judgments). As such, this body of work has also failed to lead to significant progress in the PA literature (Banks & Murphy, 1985; Ilgen et al., 1993; Longenecker et al., 1987).

The Role of Rater Motivation

The lack of notable advancement in improving PA accuracy (and strengthening the relationship between job performance and ratings of performance) is due, in part, to a focus on ability factors and assumptions that raters are motivated to provide accurate performance ratings. This notion, however, has been questioned by researchers (e.g., Cleveland & Murphy, 1992; Harris, 1994; Murphy et al., 2003; Randall & Sharples, 2010). While scholars (e.g., Longenecker et al., 1987; Mohrman & Lawler, 1983) have called for attention to the influence of rater motivation on PA outcomes in the past, only recently have researchers (e.g., Spence & Keeping, 2010; Wong & Kwong, 2007) begun to address those calls. In fact, performance rating is now being treated as a motivated behavior (Harris, 1994; Spence & Keeping, 2011). Specifically, the rater's willingness to provide accurate performance ratings is now being taken into consideration, as it has been suggested that raters make strategic decisions about what performance ratings should be recorded (e.g., Murphy et al., 2003). Therefore, inaccurate ratings

appear to be more the result of raters' lack of motivation, or unwillingness, to provide accurate performance ratings rather than a lack of ability.

In interviews with executives and managers, Longenecker and colleagues (1987) found that raters manipulate ratings when evaluating employee performance to accomplish personal goals. This is because raters are subject to a variety of pressures that motivate them to distort performance ratings (Spence & Keeping, 2010; 2011). For example, Kozlowski et al. (1998) provide a qualitative description with empirical analyses that illustrate the competing goals raters may have. Their findings, drawn from a military sample, suggest that raters find ways to convey the necessary message for administrative decisions (e.g., promotions) while still considering the organization's policy and the ratees' reactions. Specifically, their investigation suggests that raters may provide high performance ratings for all the employees they supervise, but intentionally overreport or underreport the size of an employee's comparison group (i.e., the number of individuals in the group the ratee is being compared to for promotion). This ensures that ratees with large comparison groups appear like very strong candidates for promotion while candidates with small comparison groups do not. Manipulating the ratings in such a manner allows the rater to accomplish multiple goals.

The existence of multiple goals related to performance appraisal triggers the need for the rater to determine what rating behavior is best, as the rater considers which goals to emphasize and ignore when providing performance evaluations (Cleveland & Murphy, 1992). The rater's PA goals may be in conflict, such that satisfying one goal may prevent the rater from accomplishing another goal. For example, a rater may be motivated to provide low and deflated performance ratings for an employee as part of a future termination procedure (Longenecker et al., 1987). At the same time, there may be a group norm that suggests only average and above

average ratings are acceptable (Cleveland & Murphy, 1992). Accordingly, Cleveland and Murphy (1992; Murphy & Cleveland, 1991; 1995) have proposed that rater goals play an important role in the accuracy of performance ratings. This is consistent with propositions that raters consider various outcomes (e.g., pay and promotion, other organizational members' approval of ratings) associated with the performance ratings they provide (e.g., Kane, 1994; Mohrman & Lawler, 1983).

Murphy (2008) has suggested that rater goals serve as an explanation for the weak relationship between job performance and ratings of job performance. Recent performance appraisal research (e.g., Murphy et al., 2004; Wang, Wong, & Kwong, 2010; Wong & Kwong, 2007) has found that raters have goals and provide performance ratings consistent with their goals. For example, using teaching evaluations for university instructors, Murphy and colleagues (2004) found that students' ratings of goal importance at the beginning of the semester predicted their instructor ratings at the end of the semester. Specifically, students reported the extent to which they pursued each of the following four goals: identifying instructor weaknesses, identifying instructor strengths, providing fair ratings based on performance, and motivating instructors. The results suggest that pursuing different goals results in different ratings. Similarly, Wong and Kwong (2007) found that rater goals affect performance ratings. Their findings suggest that goals affect both the accuracy and discriminability of peer evaluations. For example, when raters pursue a harmony goal (instead of a goal of identifying ratee strengths and weaknesses), ratings are inflated and there is less discriminability (i.e., less distinction between the ratees).

Instead of only focusing on peers, or students rating instructors, Wang and colleagues (2010) examined how rater goals and ratee performance levels affect the accuracy and

discriminability of supervisory performance ratings. Their results suggest that rater goals affect both accuracy and discriminability, but that it differs across performance levels. For example, when raters pursue a motivation goal, performance ratings for low performers are inflated but there are no significant distortions for high or medium performers. Finally, in a policy capturing study using a sample of experienced managers, Spence and Keeping (2010) found that avoidance and self-interest goals affect performance ratings. For example, when raters had a self-interest goal (i.e., their own performance ratings depended, in part, on the ratings they provide for ratees) ratings were inflated. These studies offer evidence that raters provide performance ratings consistent with personal goals.

The Performance Appraisal Context

Performance appraisal models that take the rater's goals into consideration are also referred to as mediated models (Murphy, 2008). Mediated models of rating behavior posit that contextual factors influence the accuracy of performance ratings through their effects on rater goals (Murphy, 2008). The performance appraisal context consists of situational, or organizational and performance appraisal factors, other than a ratee's performance, that affect performance ratings (Judge & Ferris, 1993; Murphy, 2008). While not focusing on rater goals, Greguras and colleagues (2003), for example, found that performance ratings are affected by rating purpose (i.e., administrative versus developmental), such that performance ratings conducted for developmental purposes result in higher quality of performance ratings.

Given the social nature of performance appraisal (i.e., implications for careers, rewards, interpersonal relationships: Spence & Keeping, 2013) and the various constituents (e.g., the rater's supervisor) that may have goals related to those ratings, social factors are inherent in the

performance appraisal context. Research suggests that social aspects, such as others' rating behaviors and employee influence tactics affect performance ratings (Judge & Ferris, 1993; Longenecker et al., 1987; Spence & Keeping, 2010). Taken together, this suggests that raters deliberately distort performance ratings to accomplish their goals. For example, Longenecker and colleagues (1987) note that "the Machiavellian spirit still lives" (p. 183), implying raters intentionally distort performance ratings to benefit themselves. Spence and Keeping's (2010) arguments also suggest that raters provide performance ratings that are to their benefit. For example, they argue that raters will provide higher performance ratings when their own outcomes depend on the ratings they provide. Additionally, drawing on the work by Mohrman and Lawler (1983), they agree that even if the organization encourages developmental performance ratings, raters will not be motivated to provide these when in competitive organizations because organizational norms suggest that they would not be reasonable. The social context of performance appraisal should therefore have an effect on rater goals and rating accuracy.

The social context of PA represents the social aspects (e.g., relationships between organizational members) inherent in PA (Levy & Williams, 2004). While the role of the social context has been acknowledged, it has largely been neglected in the PA literature (Cleveland & Murphy, 1992; Levy & Williams, 2004; Murphy & Cleveland, 1995). This is especially true of empirical studies. The limited empirical research on how the social context affects rater motivation and rating accuracy impedes our holistic understanding of factors influencing the rating process. It also leaves us without tests of theories proposing the social context affects performance ratings, and ultimately prevents us from improving PA accuracy.

Researchers (e.g., Cleveland & Murphy, 1992) have speculated that raters may change their goals to conform to the PA context and provide ratings consistent with those goals. While not directly focusing on rater goals or motivation, some preliminary evidence suggests that social contextual factors affect performance ratings. Specifically, Spence and Keeping (2010) found that factors such as organizational norms pertaining to performance evaluation affect ratings, even when performance is objectively presented. Kozlowski and colleagues (1998) also provide evidence for the effects of the social context on performance ratings by demonstrating that raters distort performance ratings to adapt to the performance appraisal environment. Specifically, Kozlowski and colleagues (1998) show that raters adapt to the PA context by providing ratings that are high to satisfy ratees, but manipulate the size of the comparison group to send a message about who should be promoted. Finally, Mero et al (2007) have shown that performance ratings are affected by other individuals who the rater is accountable to, even when their rating preferences are unknown. In an experiment with MBA students, they found that when raters were accountable to ratees, ratings were higher than when raters were accountable to the session administrator.

These studies illustrate the importance of some of the social contextual factors that affect performance appraisal. The social context is a broad concept that subsumes many aspects (Ferris, Jude, Rowland, & Fitzgibbons, 1994). Mitchell (1983) identified a variety of social contextual factors that affect performance appraisals based on existing research. These include interdependence between the subordinate group, similarity between the rater and ratee, power of the rater relative to the ratee, and relative comparisons among ratees. These are similar in nature to the contextual factors Mohrman and Lawler (1983) referred to as the “less easily defined aspects, such as the climate, the culture, and the nature of interpersonal relationships” (p. 178).

Levy and Williams (2004) also identified social contextual factors that affect the performance appraisal process. These include, for example, factors such as impression management, group characteristics, rater accountability, and rewards (Levy & Williams, 2004). There is no all-encompassing typology of social contextual factors that may affect the PA process. The ones that have been identified (e.g., Levy & Williams, 2004) were based on extant research findings.

For the purposes of the present study, I use existing mappings of the social context and draw on the social influence literature to provide a definition of the social context of performance appraisal and an explanation of what constitutes the construct space. The performance appraisal context represents a variety of factors, including aspects such as the legal system and the organizational culture or climate (Levy & Williams, 2004; Murphy & Cleveland, 1995). Previous performance appraisal research suggests that the social context of performance appraisal includes other individuals, or social intricacies, inherent in performance appraisal, capturing the social environment of the workplace (Levy & Williams, 2004). The social context is therefore part of the performance appraisal environment, but involves other individuals or groups, and their actions, rather than other contextual factors (e.g., rating purpose).

The individuals in the PA context and their actions constitute social influences. Generally, social influences are those that lead to changes in an individual as a result of another person or group (Raven, 1965). Extending this to performance appraisal, and drawing on the previous mappings of the social context of performance appraisal, social contextual factors should include other individuals and groups that are part of the performance appraisal context, who may affect the performance ratings rendered by the rater. These individuals, their actions, and the culture and climate they interact to create, all constitute the social context of performance appraisal. This view of the social context explicitly focuses on others in the PA context, and

retains previously identified factors such as climate, power differences, and impression management (Levy & Williams, 2004; Mohrman & Lawler, 1983; Mitchell, 1983).

Currently, there is a growing interest in how social contextual factors, especially multiple interpersonal relationships in the performance appraisal context, make performance rating more complex and simultaneously affect performance ratings (e.g., Ferris, Munyon, Basik, & Buckley, 2008; Levy & Williams, 2004). Much of the research to date, however, has focused on single goals (Cleveland & Murphy, 1992) with simplistic conceptualizations of the performance appraisal context. Kozlowski and colleagues (1998) note that these narrow conceptualizations of the performance appraisal context include factors such as appraisal purpose and rater accountability. Focusing on factors such as rating purpose neglects the social factors inherent in performance appraisal (Levy & Williams, 2004). This is problematic as social factors, such as information not related to performance (e.g., others' rating behaviors), influence performance ratings (Spence & Keeping, 2010). Yet we do not have a clear understanding of how key social contextual factors simultaneously affect the rater's motivation to provide accurate versus distorted performance ratings. It is important to understand the effect of these key social contextual influences, however, in order to be able to design effective interventions and improve PA accuracy (Spence & Keeping, 2010).

Key Social Contextual Factors in PA

The rater goal perspective (e.g., Cleveland & Murphy, 1992) proposes that the rater's goals in the PA context revolve around attaining valued outcomes or accomplishing personal goals. Many of the rater's goals should be affected by the social context of performance appraisal as these goals involve other individuals. Specifically, numerous rater goals revolve around

receiving approval from the rater's supervisor and/or subordinate group. For example, raters are motivated to provide ratings that will (1) make them appear competent to their supervisors, (2) motivate their subordinate group to maintain or improve performance, (3) maintain group harmony, and (4) avoid conflict (e.g., Cleveland & Murphy, 1992; Longenecker et al., 1987; Villanova & Bernardin, 1989; Wang et al., 2010). In one study, Wong and Kwong (2007) found that when raters were pursuing a harmony goal (i.e., maintaining positive interpersonal relationships between group members), they provided inflated performance ratings for their peers. Longenecker and colleagues' (1987) respondents reported inflating their employees' performance ratings to appear more competent to their supervisors. This is consistent with Villanova and Bernardin's (1989) conceptualization of performance rating as impression management behavior, where raters provide evaluations that will make them look good to both the supervisor and subordinate group.

Taken together, the existing performance appraisal research leads to the inference that the rater's supervisor and subordinate group are key social contextual factors in performance appraisal as the rater considers both the supervisor's and subordinate group's approval when rating performance. As performance ratings affect the allocation of valued and scarce resources, raters are likely to encounter influences from other individuals (e.g., Pfeffer et al., 1976; Zalesny, 1990). Because the rater's supervisor and subordinate group represent those the rater has the most contact with, they are likely to be the key individuals the rater considers when rating performance. A conceptual model of the antecedents of rater motivation and rating accuracy are included in Figure 1 in Appendix A. To provide a holistic picture, factors that have been shown to influence performance ratings (i.e., ratee performance and rater ability) are also depicted. This study is about the effects of the social context on rater motivation and rating accuracy. Therefore,

the effects of the rater's supervisor and subordinate group, who represent key social contextual factors, are examined in the present study.

The rater's supervisor is a salient force in the PA context for at least two reasons. First, the supervisor affects the rater's perception of what rating behaviors are appropriate and effective (Mohrman & Lawler, 1983). Specifically, a rater's motivation to provide accurate performance ratings is affected by how seriously the rater's supervisor takes performance appraisals (e.g., Longenecker et al., 1987). Respondents in Longenecker et al.'s (1987) study reported the following:

“I've learned how not to conduct the review from the bosses... but you do learn from your boss how much or what you can get away with in rating your people ... It seems that if the manager's boss takes it [the appraisal] seriously, the subordinate [manager] is more likely to follow. If the boss plays games with the review, it seems like the subordinate [manager] is more likely to do so.”
(Longenecker et al., 1987: p. 186).

This suggests that if the rater's supervisor does not take the process seriously and inflates employees' ratings, the rater is likely to infer that the performance appraisal process is not taken seriously by higher-ups in the organization and will engage in similar rating behaviors (Cleveland & Murphy, 1992; Longenecker et al., 1987). These ideas and findings in the performance appraisal literature are consistent with social learning theory (Bandura, 1977). Social learning theory, which focuses on how individuals learn, posits that this happens because learning occurs by observing others' behaviors and the resulting consequences. For example, the rater may observe that his or her supervisor does not take performance rating seriously and provides inflated performance ratings. The rater may also notice that rating inflation helps maintain group harmony and avoids damaging interpersonal relationships with other organizational members. According to social learning theory, in this situation, the rater has

learned that inflated performance ratings result in positive outcomes for the rater by observing others' rating behaviors and the consequences of those behaviors.

There is more to the supervisor's influence in the PA context than just vicarious learning through the supervisor's actions. The supervisor, due to his or her reward power, has the ability to determine many of the rater's outcomes (French & Raven, 1959; Lawler, 1973), making him or her a salient force in the PA context. For example, the rater's own outcomes may be, in part, based on the performance appraisals the rater gives to his or her employees. Additionally, the supervisor may have coercive power (French & Raven, 1959), and therefore the ability to provide the rater with undesirable outcomes. It is therefore not surprising that raters may view conducting performance appraisals as impression management, and try to appear competent to their supervisors (Villanova & Bernardin, 1989) to secure desirable outcomes and avoid undesirable ones. Ultimately, the power differential between the supervisor and the rater constitutes another important reason the rater's supervisor should be a salient source of influence in the PA context.

Another salient influence in the PA context is the rater's subordinate group, or the group of employees the rater oversees. Raters may use performance appraisal (e.g., Cleveland et al., 1989) to motivate employees to perform their jobs efficiently. Therefore, raters are likely to consider the effects performance ratings have on the members of the subordinate group, in part to ensure that they remain motivated. Raters may also consider how performance ratings could affect group dynamics, making the subordinate group a salient source of information in the performance appraisal context. For example, information about an employee taking away from group harmony may be taken into account as the performance ratings for this employee could affect other group members. Furthermore, we know that individuals are concerned with creating

and maintaining positive social relationships with other individuals and groups (Sussmann & Vecchio, 1982). This need for affiliation motivates them to maintain social relationships with others by engaging in behaviors that are approved by other members of the group (Cialdini & Goldstein, 2004). For example, Crandall, Eshleman, & O'Brien (2002) found that public expression of prejudice was related to the social group's approval of that expression. While the rater oversees the subordinate group, the subordinate group may nonetheless constitute a relevant social group for the rater. This may be because raters typically interact with the subordinate group regularly and therefore provide ratings that ensure positive interpersonal relations (e.g., Longenecker et al., 1987). In fact, performance appraisal research suggests that maintaining a positive subordinate group climate is one of the most commonly mentioned rater goals (Cleveland & Murphy, 1992). Taken together, the rater's subordinate group should be a salient source of information in the social context of PA.

The rater's supervisor and ratee's subordinate group may provide incentives for the rater to render performance ratings consistent with their goals (e.g., high performance ratings). For example, the subordinate group may provide incentives for the rater to inflate performance ratings by having a group norm for what constitutes acceptable ratings (Cleveland & Murphy, 1992). If it is communicated to the rater that average and above average ratings are the only ratings accepted by the subordinate group, then the rater should be motivated to provide inflated ratings for a poor performer, as this would satisfy the group's norms. The incentive in this case may be maintaining positive interpersonal relationships with members of the subordinate group by not violating group norms. Alternatively, the subordinate group could motivate the rater to provide deflated ratings for an employee. If the rater is aware that a ratee takes away from group

harmony, for example, the rater may be motivated to deflate that employee's ratings as part of a dismissal procedure (e.g., Longenecker et al., 1987).

Supervisors could also convey information that motivates raters to distort performance ratings (Kane, 1994). For example, if a rater perceives that his or her supervisor has a preference for high or low performance ratings for a specific ratee, the rater will be less motivated to provide accurate performance ratings (Harris, 1994). This is because some of the rater's outcomes (e.g., promotions and pay raises) are determined by the rater's supervisor (Lawler, 1973). Alternatively, if the rater knows the supervisor prefers accurate performance ratings, the rater will be motivated to provide accurate ratings. The supervisor's power (e.g., reward power: French & Raven, 1959) motivates the rater to ensure the supervisor approves of the performance ratings. Alternatively, the rater's supervisor could provide incentives for the rater to inflate performance ratings. This could be done by basing some part of the rater's monetary outcomes on the performance of the rater's subordinate (Larson, 1986).

Different individuals and groups are likely to have different goals they would like the rater to consider when providing performance evaluations (Cleveland & Murphy, 1992). Therefore, when performance appraisal is conducted in organizational settings, the rater's supervisor and subordinate group may simultaneously convey information intended to ensure that their goals are taken into account (Cleveland & Murphy, 1992). The simultaneous influence of both forces presents a complex process. In the current study, expectancy theory and social impact theory are used to clarify how supervisor and group influences combine to motivate behavior and affect rating accuracy. Expectancy theory is suitable for examining rater motivation in the PA context because it allows for the combination of numerous variables in a meaningful way (Porter & Lawler, 1968), such that, a number of possible behaviors associated with different

outcomes can be simultaneously examined. For example, Mohrman and Lawler (1983) have used expectancy theory to understand rater motivation and appraisal behavior. They considered various outcomes, including reactions of others to the ratings, interpersonal reactions of the ratee, and ability to fire the ratee, on performance appraisal behaviors (e.g., biasing performance ratings).

Expectancy theory does not, however, inform our understanding of the valences associated with the various outcomes (i.e., supervisor and group approval of performance ratings). Expectancy theory may therefore not explain the relative importance of information from the rater's supervisor and subordinate group. To understand the importance of each on rating accuracy, and what happens when it is presented simultaneously, I draw on social impact theory (Latané, 1981). Social impact theory is useful for understanding the magnitude of impact of each source of influence because it explains the magnitude of effects from sources relatively higher and lower in power than the target (Latané & Wolf, 1981). For example, Westaby and Lowe (2005) have effectively used social impact theory to examine supervisor and coworker influences on employees' safety behaviors. As such, social impact theory is integrated with expectancy theory to understand the relative effects of the information from the rater's supervisor and group on the rater.

Theoretical Frameworks

Expectancy Theory

Expectancy theory posits that motivation to perform a certain behavior is determined by two variables (Lawler, 1971). The first is the individual's effort to performance (i.e., $E \rightarrow P$) expectancy, or the extent to which one perceives that a given amount of effort will result in some

desired level of performance. $E \rightarrow P$ expectancies theoretically range from 0 to +1 (Lawler, 1973), where values close to 1 indicate the individual has a high subjective probability estimate that effort will result in performance. Values close to 0 suggest that the individual perceives that a given level of effort is unlikely to result in some desired level of performance. In the performance appraisal context, if raters have low $E \rightarrow P$ expectancies, organizations can improve a rater's $E \rightarrow P$ expectancies by providing rater training. Rater training focuses on developing skills the rater needs to observe and evaluate behavior (Murphy & Cleveland, 1995) and should therefore improve raters' $E \rightarrow P$ expectancies.

The second component is the “combination of beliefs about what the outcomes of accomplishing the intended level of performance will be and the valence of these outcomes” (Lawler, 1971: p. 108). This component consists of two parts. First, there is the performance to outcome (i.e., $P \rightarrow O$) expectancy, also referred to as instrumentality. This reflects the individual's subjective evaluation that a given level of performance will result in some outcomes. $P \rightarrow O$ expectancies also range from 0 to +1, where values close to 1 suggest a greater likelihood of an outcome being followed by a certain behavior. Values close to 0 suggest that the individual's subjective probability estimate that some level of performance will lead to a specific outcome is low. The $P \rightarrow O$ expectancy is multiplied by the valence for a given outcome. The valence (V) indicates the attractiveness of an outcome. Valences range from -1 to +1. Attractive outcomes have values close to +1, unattractive outcomes have values close to -1. Outcomes an individual finds neither attractive nor unattractive are close to 0.

Together, these two components predict the individual's motivation to engage in a certain behavior, also referred to as the motivational force (MF). The motivational force is represented by the equation: $MF = E \rightarrow P \times \sum[(P \rightarrow O)(V)]$. The summation sign illustrates that there are a

number of possible outcomes associated with any behavior and all relevant outcomes are considered. The relationship between these components is theorized to be multiplicative. Therefore, if either of the two components are 0, the motivational force will also be 0. As such, to be motivated to engage in a given behavior, an individual needs to value the outcomes associated with the behavior, perceive that behavior leads to valued outcomes, and believe that trying to engage in a behavior will result in the behavior.

Meta-analytic results indicate that expectancy theory is an effective framework for predicting work motivation and subsequent behavior (e.g., Van Eerde & Thierry, 1996). The theory has been tested in various contexts. For example, Mathieu, Tannenbaum, and Salas (1992) found support for expectancy theory predictions in a study that examined the effects of individual and situational factors on training-related motivation and effectiveness. Similarly, Jacobson and Eran (1980) effectively used expectancy theory to predict older physicians preferences for retiring versus remaining employed. Though expectancy theory has generally received support in the literature, some issues (e.g., between- versus within-subjects designs) have been debated (e.g., Van Eerde & Thierry, 1996). Specifically, Van Eerde and Thierry (1996) argue that Vroom's model was not originally intended to predict who will perform well, which would require between-subjects designs. Rather, it was intended to predict which behavioral choice an individual will make, necessitating the use of within-subjects designs. Research, however, suggests that both within- and between-subject designs are appropriate (Pinder, 1998).

Ultimately, expectancy theory is an established, well accepted, and valid theory of motivation (e.g., Miner, 1984; Sanchez et al., 2000). It is useful for the present study as it allows treatment of performance rating as a deliberate behavior, where the rater knows the actual

performance rating but manipulates it to accomplish different goals. It also permits integration of various, and potentially conflicting courses of action, and outcomes associated with each, in a meaningful way (Porter & Lawler, 1968). Expectancy theory also allows a more fine-grained examination of how contextual factors influence different motivational components and how each affects rating accuracy. As previously noted, this diagnostic approach is useful for practical recommendations (Sanchez et al., 2000), as specific motivational components that affect outcomes can be isolated.

Integrating Expectancy Theory with the Social Context of PA

Expectancy theory would suggest that raters consider three possible rating behaviors and the outcomes associated with each. In expectancy terms, the rater is likely to have some motivational force to provide accurate and distorted (i.e., inflated or deflated) performance ratings. This can be illustrated by the following three equations. Equation A represents the motivational force to provide accurate performance ratings, while equations B and C represent the motivational forces to provide inflated and deflated performance ratings, respectively. Expectancy theory suggests that raters will engage in the behavior that has the greatest motivational force (MF).

$$A) \text{ MF}_{\text{accurate}} = E_{\text{accurate}} \rightarrow P_{\text{accurate}} \times \sum[(P_{\text{accurate}} \rightarrow O) (V)]$$

$$B) \text{ MF}_{\text{inflate}} = E_{\text{inflate}} \rightarrow P_{\text{inflate}} \times \sum[(P_{\text{inflate}} \rightarrow O) (V)]$$

$$C) \text{ MF}_{\text{deflate}} = E_{\text{deflate}} \rightarrow P_{\text{deflate}} \times \sum[(P_{\text{deflate}} \rightarrow O) (V)]$$

The rater's motivational force to engage in any of the three rating behaviors is determined by the rater's $\sum[(P \rightarrow O) (V)]$ and $E \rightarrow P$ expectancies. It is important to clarify the treatment of

$E \rightarrow P$ expectancies in the present study before addressing the valences of outcomes raters are likely to consider when providing performance ratings and contextual effects on $P \rightarrow O$ expectancies. The rater's $E_{\text{accurate}} \rightarrow P_{\text{accurate}}$ expectancy, for example, indicates the extent to which the rater believes that trying to report accurate performance ratings will result in providing accurate performance ratings. In interviews with executives and managers, Longenecker and colleagues (1987) found that although raters reported intentionally distorting performance ratings, no raters reported feeling a lack of ability to rate accurately. With proper training, resources, and ability, individuals should perceive that they can accurately rate performance if they try to do so. With inflated and deflated performance ratings, raters should have high subjective probability estimates that trying to inflate or deflate performance ratings should result in inflated and deflated performance ratings, respectively. Additionally, $E \rightarrow P$ expectancies are not influenced by contextual factors. *Ceteris paribus*, $E \rightarrow P$ expectancies for accurate, inflated, and deflated performance ratings should be high and are assumed to be equal to 1 in the hypotheses developed in this study.

$P \rightarrow O$ expectancies are most strongly affected by the actual situation as individuals look to aspects of the situation (e.g., organizational pay and promotion systems and leaders' behaviors) to determine how specific behaviors result in different outcomes (Lawler, 1973). It can therefore be argued that the performance appraisal context will affect $P \rightarrow O$ expectancies. Stated differently, the PA context affects the extent to which outcomes (e.g., promotions and positive interpersonal relationships) are a result of accurate, inflated, or deflated performance ratings. If we know how a person perceives the situation, as well as values the individual assigns to various outcomes, we can predict behavior (Lawler, 1973). It is important to note that

expectancy theory suggests that individuals have limited rationality and do not consider all possible alternatives, only the most salient potential outcomes (Lawler, 1971).

The possible outcomes resulting from a given behavior are valued to the extent that they satisfy the individual's needs, and as all outcomes do not equally satisfy needs, they are not equally valued (Lawler & Porter, 1967). Raters may have various needs that can be fulfilled in the workplace. One need theory, focusing on the work context, developed by McClelland and colleagues (e.g., McClelland, 1961), posits that individuals have needs for achievement, power, and affiliation. Extending another needs theory (i.e., Maslow's hierarchy of needs) to the work domain, Hall and Nougaim (1968) found support for these needs. For example, they found that managers' needs for affiliation and achievement increased as managers advanced in the organization. Research suggests that individuals have a need for affiliation, or belonging, which can be satisfied in the workplace (Wiesenfeld, Reghuram, & Garud, 2001). As the need for affiliation represents a desire for close interpersonal relationships (Hall & Nougaim, 1968), both the rater's supervisor and subordinate group should be able to satisfy this need. The need for achievement and power can also be satisfied in the workplace, through pay raises and promotions, for example. Mitchell and Mickel (1999) suggest that money is important to individuals, in part due to its symbolic value that suggests achievement. The need for power may be satisfied with promotions which allow the rater to move up the organizational hierarchy and acquire greater power. Therefore, the rater's supervisor may also be able to satisfy the needs for power and achievement, given the supervisor's reward power, or ability to administer rewards (French & Raven, 1959). Considering the supervisor's and subordinate group's ability to satisfy these needs, supervisor and subordinate group approval will be valued outcomes for the rater.

In the current study, the focus is on how information from the rater's supervisor (i.e., expressed rating preference) and subordinate group (i.e., opinion about the ratee) affects the rater's $P \rightarrow O$ expectancies². Information from the rater's supervisor and group will be most salient, as these parties are the most immediate in the social situation, or social context, in which performance appraisal is embedded (Levy & Williams, 2004). Moreover, because performance appraisal is an inherently social process that involves various constituents (Judge & Ferris, 1993; Kozlowski et al., 1998; Murphy & Cleveland, 1993), it is likely that the rater's supervisor and/or subordinate group will provide information that influences raters (e.g., Longenecker et al., 1987). When this occurs, the supervisor's and subordinate group's approval of the ratings are two outcomes the rater will consider. The supervisor's preference and subordinate group's opinion are, therefore, likely to affect the raters' $P \rightarrow O$ expectancies. Specifically, the extent to which accurate, inflated, or deflated performance ratings result in group and supervisor approval depends on the supervisor's preference regarding performance ratings and group's opinion about the ratee. For example, when the rater's supervisor provides information that encourages rating inflation, the rater's $P \rightarrow O$ expectancy for inflated ratings resulting in supervisor approval of ratings will be high.

The supervisor and subordinate group will also affect the possible outcomes associated with different ratings. If neither the subordinate group nor the supervisor provide information

² Other factors, such as ratees' and top managers' information may also affect performance ratings. Considerable research suggests that ratees' behaviors, such as impression management (Harris, Zivnuska, Kacmar, & Shaw, 2007) or subordinate influence tactics (Ferris et al., 1994) affect performance ratings. This area has been given considerable research attention, and is therefore not the focus of the present study. Senior, or top managers, could also have an effect on performance ratings. This effect, however, should not be as strong as the effect of the rater's immediate supervisor and subordinate group. Turner, Chmiel, Hershcovis, and Walls (2010), for example, found that senior managers are not as influential as they are not as salient or proximal to most individuals in the organization. Therefore, the focus of the current study is on the effects of information from the rater's supervisor and subordinate group on rating accuracy.

regarding performance ratings, expectancy theory would suggest that the rater is unlikely to perceive group or supervisor approval of the ratings as an outcome. This is because only a limited number of outcomes are considered. When the supervisor or group provide information relevant to performance ratings, the approval of each is likely to become a salient outcome. For example, if the rater's supervisor encourages rating inflation, the rater will expect inflated ratings to result in approval from the supervisor. This is an outcome the rater should value given the supervisor's ability to administer positive and negative outcomes. Salient supervisor preferences make supervisor approval of ratings a possible outcome with a positive valence. When the subordinate group provides information regarding the ratee, the subordinate group's approval will also be an outcome that is salient to the rater. The group's approval should also have a positive valence as the group can provide benefits such as social acceptance (e.g., Lawler, 1973). Expectancy theory suggests that the behavior that results in the most valued outcomes is the behavior the rater will engage in. To understand the valences associated with approval from the rater's supervisor and subordinate group, and the relative importance of each, I draw on social impact theory.

Social Impact Theory

The present study focuses on the effects of the rater's supervisor and subordinate group on rater motivation and performance ratings. To understand their simultaneous influence, social impact theory is used. Social impact theory is a broad theory, with specific predictions, about the effect, or influence, of social actors on individuals and groups (e.g., Latané & Nida, 1980). These effects are referred to as social impact (e.g., Latané, 1981). Social impact has been defined as “any of the great variety of changes in physiological states and subjective feelings, motives and emotions, cognitions and beliefs, values and behavior, that occur in an individual human or

animal as a result of the real, implied, or imagined presence or actions of other individuals” (Latané & Nida, 1980: p. 5). Social impact theory postulates that social impact, or the extent to which an individual is influenced by others, is a multiplicative function of the strength, immediacy, and number of the source (i.e., the individual or group) exerting social influence (Latané, 1981).

Strength refers to the power or social status of the source of influence (Sedikides & Jackson, 1990). In a review of social impact theory, Mullen (1985) noted that strength has generally involved some aspect of power, status, or prestige. Evidence suggests that source strength has an effect on social impact. Sedikides and Jackson (1990), for example, found that a greater percentage of individuals complied with the request when the individual making it was uniformed, and therefore appeared legitimate, than when the individual was casually dressed. Williams and Williams’ (1989) study found that source strength is especially important when the target of influence is concerned with, or motivated by the source of the impact, rather than focused on internal motivations.

The second component, immediacy, refers to the “closeness in space or time and absence of intervening barriers or filters” (Latané & Nida, 1980: 7). Sedikides and Jackson (1990) note that immediacy can be either the physical or psychological distance. Generally, immediacy has been treated as an indicator of interpersonal distance (Mullen, 1985). For example, immediacy has been operationalized as the size of an image, geographical location, distance from the door, or physical distance between individuals (Argo, Dahl, & Manchanda, 2005; Jackson & Latané, 1981, Latané & Wolf, 1981). Research has found that immediacy has an effect on compliance. For example, Sedikides and Jackson (1990) found that a higher percentage of individuals were more compliant right after a request was made than after some time elapsed. Similarly, Argo et

al. (2005) found that greater immediacy to others had a greater effect on purchasing decisions. Theoretically, the greater the immediacy of the source of influence to the target, the greater the social impact the target experiences (Latané & Wolfe, 1981). One reason greater immediacy is related to greater impact may be due to a greater ability to monitor the target when immediacy is high versus low (Sedikides & Jackson, 1990).

The third component, number of the source, simply refers to the number of individuals present who constitute the source of influence (e.g., Latané, 1981). According to social impact theory, the first source (i.e., the first individual exerting influence) will have the strongest effect on the target and each subsequent source will increase impact, but only slightly (Sedikides & Jackson, 1990). For example, the number of people present has marginally decreasing impact, such that the difference between 0 and 1 is greater than the difference between 99 and 100 (Latané & Nida, 1980). Latané (1981) arrived at the idea of marginal impact based on his review of studies on influence. One of these studies was an experiment conducted by Darley and Latané (1968) that illustrated an exponential decrease in helping as a function of the number of bystanders present. While social impact theory does not explicate the mechanisms underlying number, individuals' needs for affiliation and accuracy may aid in explaining why this happens. One of the reasons individuals are affected by social influences, and act in accordance with others' expectations, is because they want to be accepted and belong (Cialdini & Goldstein, 2004). Individuals are also susceptible to social influence as the information that is provided by others may provide relevant insights about reality (Deutsch & Gerard, 1954). Compared to when no information is available, the first source of influence is likely to have the greatest impact as novel and relevant information about reality is encountered. Each subsequent source should have some impact, as the need for accuracy, for example, is satisfied to a greater extent when

numerous individuals share the same information compared to only one person. Each subsequent individual providing the same information should have diminishing effects, however, as the information is no longer novel and agreement has been established.

Social impact theory has been criticized for failing to explain why effects occur (Mullen, 1985). It is important to note, however, that social impact theory was not intended to explain the underlying mechanisms of social influence (Jackson, 1986). Rather, social impact theory predicts the magnitude of influence and is most useful when combined with other theories that explicate the mechanisms underlying social influence (Latané & Nida, 1980). For example, drawing on French and Raven's work on power, Sedikides and Jackson (1990) have suggested that strength may have an impact because the source of influence in their study (i.e., a uniformed zoo employee) was perceived as having legitimate authority to make a request. As such, social impact theory should be used in conjunction with other theories that are relevant to the area under investigation (Latané & Wolfe, 1981).

Integrating Social Impact Theory with the Social Context of PA

To apply social impact theory to the social context of performance appraisal, each of the three components (i.e., number, strength, and immediacy) determining impact is discussed in the context of performance appraisal. Consistent with previous conceptualizations of number (Mullen, 1985), the number of the source is treated as the number of individuals constituting the source of influence. In the performance appraisal context, for the supervisor, the number of the source would be equal to one as the supervisor represents a single individual. For the subordinate group, the number of the source would be greater than one (i.e., two or more individuals), given the definitions of a group. Specifically, the Merriam-Webster dictionary defines a group as “a

number of individuals assembled together or having some unifying relationship” or “two or more figures forming a complete unit in a composition.” Therefore, in the present study, the number of the source is greater for the subordinate group than the supervisor³.

In line with the literature, strength for the supervisor and the subordinate group is conceptualized as the power of the source relative to the target (e.g., Sedikides & Jackson, 1990; Latané, 1981). In the performance appraisal context, the supervisor will be relatively more powerful than the target (i.e., rater) for a number of reasons. First, due to the supervisor’s relatively higher position in the organizational hierarchy, the supervisor may be perceived to have a legitimate right to influence behavior (French & Raven, 1959; Magee & Galinsky, 2008). The hierarchical position the supervisor occupies should also provide the supervisor with reward and coercive power (French & Raven, 1959). Specifically, the supervisor is generally able to administer rewards such as pay raises, or punishments such as undesirable work assignments. Therefore, the supervisor should be relatively more powerful than the target. The rater’s subordinate group, which the rater oversees, is hierarchically lower than the supervisor. Social impact theory, however, suggests that sources lower in power than the target can still influence the target. In the performance appraisal context, this may be because the subordinate group can provide rewards more social in nature, such as social approval (Lawler, 1973), which ensures positive interpersonal interactions. Taken together, the group has some power relative to the target, but the supervisor has more power (i.e., strength) than the target and the group.

³ While social impact theory postulates that the number of the source pertains to the number of individuals comprising the source of influence (e.g., Latané, 1981), it is possible it also operates through the number of influence attempts. Given that the subordinate group consists of multiple individuals who provide consistent information, it is as likely that the influence attempts from the subordinate group are greater in number than those from the supervisor. This interpretation would also lead to the inference that the number is greater for the subordinate group than the supervisor.

The final component, immediacy, has generally been conceptualized as interpersonal distance (e.g., Sedikides & Jackson, 1990). However, it can also refer to the absence of filters or intervening barriers between the source and the target of the influence (Latané, 1981). Latané and Nida (1980) note that social impact theory is intended to apply to a broad range of situations, and may therefore need simple modifications when applied to different contexts. Specifically, Latané and Nida (1980) note that “any situation in which social impact occurs is representable within this basic three-unit fabrication or some simple modification of it” (Latané & Nida, 1980: p. 28). One of the reasons immediacy predicts social impact is because closeness (or access) increases the likelihood of evaluation of the target by the source (Sedikides & Jackson, 1990). Therefore, to fit the performance appraisal context, immediacy is conceptualized here as the absence of intervening barriers between the source of the influence and access to the performance ratings provided by the rater. In the performance appraisal context, the rater’s supervisor has direct access to the performance ratings provided by the rater for all ratees. The group may also be aware of group members’ performance ratings. Drawing on social comparison research (e.g., Festinger, 1954), it appears that in order to evaluate his or her performance ratings and make sense of them, a ratee would compare his/her ratings to those of other group members. This necessitates sharing performance ratings. Therefore, immediacy is not different in the present study.

Social impact theory would predict that information from the rater’s supervisor and subordinate group will have an effect on the rater as both sources should have some impact due to their strength, immediacy, or number. Providing support for social impact theory, Westaby and Lowe (2005), for example, found that both the supervisor and group affect individuals’ risk-taking orientations. Specifically, the employee’s perception that the supervisor does not approve

of employee risk-taking and the extent to which the employee perceives coworkers take risks affect the employee's own risk-taking behaviors. Therefore, when the supervisor and subordinate group provide information in the performance appraisal context, the performance evaluations the rater provides for an employee should be influenced by the rater's supervisor and subordinate group.

Social Impact Theory, Expectancy Theory, and the PA Context

Applying social impact theory to the performance appraisal context suggests that information conveyed by the supervisor should have greater impact on raters than information conveyed by the rater's subordinate group because the supervisor has greater strength relative to the subordinate group. However, both the subordinate group's and the supervisor's approval of the ratings should have some effect as both should represent valued outcomes for the rater. Integrating social impact theory with expectancy theory and the PA context suggests that the positive valence associated with supervisor approval of the ratings will be stronger than the positive valence associated with subordinate group approval of the ratings when both sources provide information. This is because the supervisor is more powerful and able to administer more positive and negative outcomes than the subordinate group. Using social impact theory and expectancy theory, the effects of different combinations of information from the rater's supervisor and subordinate group on rater motivation and performance rating accuracy are examined.

The Present Study

Overview

In this section, I first discuss the social contextual forces in the performance appraisal context that are addressed in the present study. Then, drawing on expectancy theory and the literature on social influence, I explain why information from the supervisor and subordinate group affects rater motivation and the accuracy of performance ratings, and develop hypotheses that will be tested. Subsequently, I integrate expectancy theory with social impact theory to predict the interactive effects of the supervisor and subordinate group on rater motivation and rating accuracy. Specifically, I examine what happens when both the rater's supervisor and subordinate group provide information regarding performance ratings and develop hypotheses around this.

Forces in the PA Context

As previously noted, the rater's supervisor will be salient in the performance appraisal context. The rater's supervisor may provide information that motivates the rater to inflate performance ratings. For example, the rater's supervisor could ask the rater to provide elevated performance ratings (i.e., ratings that are higher than deserved) for a specific employee. In such an instance, the supervisor could be friends with the employee and may want to ensure that his or her friend receives the benefits associated with high performance ratings (e.g., a pay raise). The rater will be motivated to provide ratings consistent with the supervisor's preferences as the supervisor can provide the rater with valued outcomes such as pay and promotions (Lawler, 1973). Pay may satisfy the rater's need for achievement, as money has symbolic value (Mitchell & Mickel, 1999). The promotion may satisfy the rater's need for power and achievement. The

rater will therefore be interested in providing ratings consistent with the supervisor's preferences to ensure the supervisor approves of the ratings and, as a result, provides the rater with these valued outcomes.

Alternatively, the rater's supervisor could ask the rater to deflate a ratee's performance ratings. For example, the supervisor may ask the rater to provide deflated (i.e., lower than deserved) performance ratings for a low performer. The supervisor may do this because he or she doesn't like the ratee or to ensure that there is a paper trail for a future dismissal of that ratee (e.g., Longenecker et al., 1987). If the rater's supervisor does not have any specific goals associated with the ratee, the supervisor may have no preference regarding that ratee's performance evaluation. In this situation, the rater's supervisor would offer no information pertaining to the ratee's performance ratings.

The rater's subordinate group may also provide incentives for the rater to provide accurate, inflated, or deflated performance ratings for a certain ratee. For example, the rater may want to maintain group harmony among his or her subordinates by inflating an average, but well-liked employee's performance ratings. The rater's subordinate group, however, is generally less instrumental than the rater's supervisor in providing formal organizational rewards such as promotions (e.g., Lawler, 1973). The rater oversees the subordinate group, and therefore occupies a higher position in the organization, such that the subordinate group does not administer formal organizational rewards to the rater. Regardless, the subordinate group can have an effect on performance ratings because the rater is motivated to maintain group harmony and positive interpersonal relationships (e.g., Cleveland & Murphy, 1992, Wong & Kwong, 2007).

For example, the subordinate group may express that a specific ratee does not contribute to, or takes away from, the contextual environment of the workplace. The contextual environment of the workplace is created through employees helping one another and being generally good employees (Van Scotter & Motowidlo, 1996). To maintain harmony within the subordinate group and between the rater and the subordinate group, the rater may be motivated to provide low and deflated ratings for that individual. While the subordinate group as a whole does not have access to performance ratings for each employee, it is possible that the members of the rater's subordinate group may infer that the employee received high or low performance ratings. Specifically, if an employee who is not well-liked by the group is promoted, the group will infer that the employee received high performance ratings which resulted in the promotion. This could strain interpersonal relationships between the rater and the subordinate group.

Alternatively, a specific ratee may be an average performer but may make significant contributions to the organizational environment, and therefore be an integral and liked member of the subordinate group. In that case, the rater may perceive that in order to avoid conflict with the group and to maintain harmony within the subordinate group, he or she should provide above average, and therefore inflated, performance ratings for that specific employee. Inflated ratings should ensure that the employee stays motivated (e.g., Longenecker et al., 1987) and continues contributing to the contextual environment, ensuring group harmony among employees. Stated differently, the subordinate group will motivate the rater to provide inflated performance ratings for that specific ratee.

Hypothesis Development and Hypotheses

Inflation versus Deflation Information

The performance appraisal literature suggests that rating inflation, or leniency, is more common. Given its higher base rate occurrence, rating inflation has been of greater interest to scholars and practitioners than rating deflation and has received considerably more research attention (Bernardin, Tyler, & Villanova, 2009). Therefore, prior to addressing the effects of social contextual factors on performance ratings, it is important to clarify two types of potential rating distortion: inflation and deflation. This initial discussion is simplified, and a more complex treatment of the type of distortion *and* source of information will follow in subsequent sections.

Performance appraisal research indicates that raters are motivated to provide high performance ratings (e.g., Longenecker et al., 1987) because many positive outcomes are associated with giving employees high ratings. For example, high performance evaluations allow the rater to secure pay raises and promotions for employees (Cleveland & Murphy, 1992). Additionally, high performance ratings ensure positive interpersonal interactions between the rater and his or her employees (e.g., Harris, 1994). Having employees with high performance evaluations also allows the rater to appear competent to his or her supervisor because these create an impression of high performing employees (Villanova & Bernardin, 1989). For the rater, few negative outcomes result from employees being awarded high performance ratings (Murphy & Cleveland, 1995).

Providing low performance ratings, however, has a number of negative consequences for the rater. For example, low performance ratings are associated with potential conflict and strained interpersonal relationships with ratees (e.g., Harris, 1994). Additionally, when an

employee has a low performance evaluation, the employee is unlikely to receive a pay raise given the limited budgets organizations have (e.g., Murphy & Cleveland, 1995). Therefore, it is not surprising managers report being hesitant to provide low performance evaluations (e.g., Longenecker et al., 1987). Given the consequences associated with providing high versus low performance ratings for the rater, high performance ratings are much more common than low performance ratings.

This has implications for rating inflation and deflation. Because positive outcomes are associated with high ratings, raters should perceive generally positive outcomes as a result of providing inflated performance ratings. On the contrary, low performance ratings tend to be associated with negative outcomes for the rater. Therefore, relatively more negative outcomes should be perceived as a resulting of providing deflated performance ratings. Research (e.g., Longenecker et al., 1987) suggests that while raters are reluctant to deflate performance ratings, when it benefits them, they will provide low and deflated performance ratings for employees. Empirical research in the performance appraisal literature has not adequately addressed the issue of rating deflation, however, given its lower base rate of occurrence. Studies have also failed to investigate whether rater motivation and performance ratings are affected to different degrees by information encouraging high versus low performance ratings. As such, we do not have sufficient evidence regarding the effects of information encouraging high versus low performance ratings on rater motivation and rating accuracy.

When both the supervisor and the subordinate group provide information encouraging high performance ratings, the outcomes associated with providing high, and in the case of an average performer, inflated performance ratings, results in largely positive outcomes for the rater. Specifically, the supervisor and subordinate group will approve of the ratings and the ratee

should also be satisfied with the higher performance ratings as evaluations are tied to valued outcomes, such as pay raises. When the rater's supervisor and subordinate group both provide information encouraging low performance ratings, however, the outcomes the rater experiences may no longer be entirely positive. If the rater provides an average performer with low, and therefore deflated performance ratings, consistent with the information from the supervisor and subordinate group, the subordinate group and supervisor should approve of the ratings. At the same time, however, the deflated performance ratings may harm the rater's relationship with the ratee and result in future conflict, and other potentially undesirable consequences (e.g., awkward feedback meetings).

Considering the outcomes associated with rendering high and low performance ratings, *ceteris paribus*, expectancy theory would posit that the rater is more motivated to provide ratings consistent with supervisor and subordinate group information when the information encourages rating inflation rather than rating deflation. This is because inflated ratings will result in largely positive outcomes while deflated performance ratings are also likely to result in some negative outcomes. Ultimately, inflated performance ratings result in more overall positive outcomes than deflated performance ratings for the rater. In expectancy theory terms, $\sum[P_{\text{inflate}} \rightarrow O(V)]$ is greater than $\sum[P_{\text{deflate}} \rightarrow O(V)]$. As such, the motivational force associated with inflating performance ratings is relatively stronger than the motivational force to deflate performance ratings. The rater will therefore be more motivated to provide ratings consistent with the subordinate group's and supervisor's information, which conveys their preferences, when the information prompts rating inflation compared to rating deflation. Extending this logic, inflation information should have a stronger effect on performance ratings than deflation information. Specifically, when both the supervisor and the subordinate group provide information

encouraging high performance ratings, there are no negative consequences preventing the rater from inflating performance ratings to the greatest extent possible. When both provide information supporting low performance ratings, however, potential negative consequences associated with the ratee may attenuate the extent to which ratings are deflated. This suggests that inflation information will have a stronger effect on performance ratings than deflation information.

Hypothesis 1: Inflation information will have a stronger effect on performance ratings than deflation information.

Main Effects of Supervisor

Researchers (e.g., Mohrman & Lawler, 1983) have proposed that the rater's supervisor affects the rater's performance appraisal perceptions. In interviews with executives and managers, Longenecker and colleagues (1987) found that the manner in which higher ups treat performance appraisals affects how managers rate performance. Specifically, if higher ups use the performance appraisal system to accomplish personal goals, managers will do the same and not be concerned with the accuracy of performance ratings. This anecdotal evidence suggests that raters are influenced by their supervisors when completing performance evaluations.

Performance appraisal research (e.g., Villanova & Bernardin, 1989) also suggests that the supervisor's approval of performance ratings is something raters consider when providing performance ratings.

The research on social influence provides a number of explanations for why the rater's supervisor may affect performance evaluations and why supervisor approval of ratings is a rater goal. For example, individuals change their behavior, or comply with the requests of those in

positions of authority, such as supervisors, for a number of reasons (e.g., Levy, Collins, & Nail, 1998). First, the supervisor is in a position of authority. Given the supervisor's position, relative to the rater, the supervisor has legitimate power (i.e., occupies a position that allows him/her to direct behavior), and therefore a right to influence behavior (French & Raven, 1959). Due to the different positions in the organizational hierarchy, the individual may feel obligated to comply with the supervisor's information (Raven, 2001). Second, the supervisor may also be perceived to have expert power (i.e., relevant expertise or knowledge: French & Raven, 1959) and information power (e.g., possesses relevant information: Raven, 2001). Therefore, information from the supervisor may be informative about how the performance appraisal system should be used and what ratings the ratee should receive. Loi, Lai, and Lam (2012), for example, found that supervisor's affective commitment was related to the employee's affective commitment. The authors reason that the supervisor's level of affective commitment provides useful information as the supervisor is a salient member of the organization and is likely to have information the employee does not possess.

The supervisor is also a source of normative social influence, providing the rater with information about what behaviors are expected and approved of (e.g., Deutsch & Gerard, 1954; Kaplan & Miller, 1987; Westaby & Lowe, 2004). This not only signals what the supervisor expects but also what behaviors are effective (Reno et al., 1993) as the supervisor determines the rater's outcomes (e.g., Stajkovic & Luthans, 2001). Finally, the rater may also feel accountable to the supervisor, and research (Mero et al., 2007) suggests accountability to higher-ups affects performance ratings. Together, the social influence and PA literatures indicate the rater's supervisor should have an effect on the rater. Thus, when the rater's supervisor provides

information pertaining to an employee's performance evaluation, it should have an effect on performance ratings.

When the supervisor provides information that conveys a preference for high or low performance ratings, the supervisor's approval will be a salient outcome. The social influence literature suggests that raters will act in accordance with the supervisor's preferences. This is consistent with the rater goal perspective (e.g., Cleveland & Murphy, 1992) that suggests the rater's goals revolve around attaining valued outcomes or accomplishing personal goals. Therefore, when the only salient goal is the supervisor's approval of ratings, raters should be motivated to provide performance ratings consistent with the supervisor's preferences, as supervisor approval will be a salient and valued outcome. When this is the case, the rater's $P \rightarrow O$ expectancies that valued outcomes result from distorted performance ratings will be high, while $P \rightarrow O$ expectancies for accurate ratings will be very low. Therefore, when the rater's supervisor provides information encouraging rating deflation (inflation), ratings should be lower (higher) than when no information is provided. Given that there should be no negative consequences associated with high performance ratings for the rater, he or she should experience less resistance to inflate performance ratings than to deflate performance ratings, as the latter behavior could result in negative outcomes associated with the ratee. Therefore, supervisor inflation information will have a stronger effect on distortion than supervisor deflation information.

Hypothesis 2a: Performance ratings will be lower when the supervisor provides deflation information than when no information is provided.

Hypothesis 2b: Performance ratings will be higher when the supervisor provides inflation information than when no information is provided.

Hypothesis 2c: Supervisor inflation information will have a stronger effect on performance ratings than supervisor deflation information.

Main Effects of Subordinate Group

The performance appraisal literature suggests that maintaining a positive subordinate group climate is one of the most commonly mentioned rater goals (Cleveland & Murphy, 1992). Research (e.g., Wong & Kwong, 2007) has also shown that when maintaining subordinate group harmony is a goal for the rater, performance ratings are distorted. As previously noted, Wong and Kwong (2007), for example, found that when raters pursue a harmony goal, peer ratings are inflated and there is less discriminability (compared to an identification of strengths and weaknesses goal). Research also suggests that raters consider the effects of performance ratings on the subordinate group's motivation (e.g., Longenecker et al., 1987; Wang et al., 2010).

The literature on social influence suggests that the group's preferences will have an effect on the rater's performance evaluations. This is because individuals conform to group pressures (Levy et al., 1998). The group can be a source of both normative and informational social influence (e.g., Westaby & Lowe, 2004). Specifically, the subordinate group can signal what behaviors are expected and approved of. The group can also provide relevant information, which may affect the individual as he or she is trying to make a decision (e.g., Raven, 1992). For example, the subordinate group may provide information about an employee the rater may not have. The social influence literature suggests that individuals are concerned with creating and maintaining positive relationships with other individuals and groups (Sussmann & Vecchio, 1982). Therefore, providing high performance ratings for an employee the group dislikes, even if the employee is a high performer, may result in damaging interpersonal relationships with members of the subordinate group.

The subordinate group's approval of the ratings should therefore have a positive and somewhat strong valence. Expectancy theory would predict that the information from the rater's subordinate group will have an effect on performance evaluations provided by the rater. When the subordinate groups' approval of the performance ratings is the most salient outcome to the rater, the rater will be motivated to provide ratings consistent with the information provided by the subordinate group as valued outcomes will be a result of rating consistently with the subordinate group's information. For example, when the subordinate group conveys information encouraging low performance ratings, the rater's $P \rightarrow O$ expectancy that valued outcomes result from low performance ratings will be high and $P \rightarrow O$ expectancy for accurate ratings resulting in valued outcomes will be low. Therefore, the motivational force to provide low performance ratings will be higher than the motivational force to provide accurate performance ratings. As with the information from the supervisor, inflation information from the group should have a stronger effect on rating distortion than deflation information from the group because inflating performance ratings is unlikely to result in any negative outcomes for the rater.

Hypothesis 3a: Performance ratings will be lower when the subordinate group provides deflation information than when no information is provided.

Hypothesis 3b: Performance ratings will be higher when the subordinate group provides inflation information than when no information is provided.

Hypothesis 3c: Subordinate group inflation information will have a stronger effect on performance ratings than subordinate group deflation information.

Combined Effects of Supervisor and Subordinate Group

So far, I have argued that information from the rater's supervisor and subordinate group will have effects on rater motivation, and ultimately, the accuracy of performance ratings

provided by the rater. In the performance appraisal context, it is likely that raters will encounter simultaneous information from both the supervisor and subordinate group (e.g., Kozlowski et al., 1998). Specifically, both the supervisor and the subordinate group could provide information that encourages high or low performance ratings. Alternatively, the supervisor and the subordinate group could provide inconsistent information, such that the supervisor information encourages high performance ratings and the subordinate group's information encourages low performance ratings, and vice versa. Therefore, the rest of the hypotheses focus on the interactive effects of the rater's supervisor and subordinate group on rater motivation and rating accuracy.

When both the rater's supervisor and subordinate group provide information encouraging either high or low performance ratings, the rater will perceive that there is consensus, or agreement regarding the direction of distortion. According to expectancy theory, $\sum[P_{\text{inflate}} \rightarrow O(V)]$ and $\sum[P_{\text{deflate}} \rightarrow O(V)]$ will be high when both the rater's supervisor and subordinate group prefer high and low performance ratings, respectively. In other words, when both the rater's supervisor and subordinate group provide information supporting high performance ratings or low performance ratings, positive outcomes associated with the subordinate group and the supervisor will result from inflated and deflated performance ratings, respectively. As there is perceived consensus, the motivational force to inflate or deflate performance ratings in such performance appraisal contexts will outweigh the motivational force to provide accurate performance ratings. This is because positive outcomes (i.e., supervisor and subordinate group approval) result from either high or low performance ratings and there are no salient outcomes associated with providing accurate performance ratings.

Therefore, when both the subordinate group and the supervisor provide information that encourages the rater to distort ratings in the same direction (i.e., inflate or deflate performance

ratings), the rater will be motivated to provide ratings consistent with the supervisor's and subordinate group's preferences. In expectancy theory terms, $\sum[P_{\text{inflate}} \rightarrow O(V)]$ will be higher when both the rater's supervisor and subordinate group prefer high performance ratings than when only one source provides information encouraging high performance ratings. Similarly, $\sum[P_{\text{deflate}} \rightarrow O(V)]$ will be higher when both the rater's supervisor and subordinate group prefer low performance ratings than when only one source provides information encouraging low performance ratings. Therefore, ratings will be higher (lower) when both sources convey information encouraging high (low) ratings than when only one source provides information encouraging high (low) ratings. Performance ratings should also be higher when both the rater's supervisor and subordinate group convey information supporting high performance ratings than when both convey information supporting low performance ratings. Furthermore, the effect of simultaneous inflation information from both sources combined should be stronger than the effect of simultaneous deflation information encountered from both sources.

Hypothesis 4: Performance ratings will be higher when both the supervisor and subordinate group provide consistent inflation information than when (a) only the supervisor provides inflation information, (b) only the subordinate group provides inflation information, or (c) no information is provided.

Hypothesis 5: Performance ratings will be lower when both the supervisor and subordinate group provide consistent deflation information than when (a) only the supervisor provides deflation information, (b) only the subordinate group provides deflation information, or (c) no information is provided.

Hypothesis 6: Performance ratings will be higher when both the supervisor and subordinate group provide consistent inflation information than deflation information.

Hypothesis 7: Subordinate group and supervisor inflation information will have a stronger effect on performance ratings than subordinate group and supervisor deflation information.

There may also be instances when both the subordinate group and supervisor provide information that motivates the rater to engage in different rating behaviors. This is consistent with research suggesting raters have multiple and sometimes conflicting goals when evaluating performance (e.g., Cleveland & Murphy, 1992; Kozlowski et al., 1998). The literature that has been reviewed illustrates that information from the supervisor and the subordinate group affects raters. This is because the supervisor's and subordinate group's approval of the ratings are attractive outcomes as they can satisfy a number of rater needs. To make predictions regarding the interactive effects of conflicting information from the rater's supervisor and group on rater motivation and performance ratings, I incorporate social impact theory.

Social impact theory postulates that when influence is exerted in different directions, compliance is reduced since the sources of influence are not in consensus (Latané & Wolf, 1981). The sources of influence are directing behavior in opposite directions, and taking away from the overall impact in either direction as the number of the source providing specific information is reduced. Specifically, compared to when both the supervisor and subordinate group convey information encouraging high performance ratings, the number of the source⁴ encouraging high performance ratings is reduced when one provides information encouraging low performance ratings. This suggests that the impact of the rater's supervisor or subordinate group is less than if consensus exists (i.e., both the rater's supervisor and subordinate group encourage the same rating behavior). This happens because the rater's motivational force to provide high or low performance ratings is lower when only one source of influence encourages high or low performance ratings than when both sources encourage the same rating behavior.

⁴ The number of the source refers to the number of individuals exerting influence in the a specific (i.e., the same) direction, or encouraging the same behavior.

For example, when the supervisor provides information encouraging low performance ratings for an employee, and the subordinate group provides information encouraging high performance ratings for that same employee, $\sum[P_{\text{deflate}} \rightarrow O(V)]$ will be lower than when the subordinate group also provides information encouraging low performance ratings. In the former case, only some positive outcomes (i.e., supervisor approval) result from low performance ratings. In the latter scenario, however, both positive outcomes (i.e., supervisor and subordinate group approval) result from low performance ratings. Therefore, I predict that the performance ratings will be higher (lower) when both the supervisor and subordinate group provide inflation (deflation) information than when they provide inconsistent information.

Hypothesis 8: Performance ratings will be higher when both the subordinate group and supervisor provide inflation information than when (a) the supervisor provides inflation information and the subordinate group provides deflation information, or (b) when the supervisor provides deflation information and the subordinate group provides inflation information.

Hypothesis 9: Performance ratings will be lower when both the subordinate group and supervisor provide deflation information than when (a) the supervisor provides inflation information and the subordinate group provides deflation information, or (b) when the supervisor provides deflation information and the subordinate group provides inflation information.

Social impact theory also postulates that when two conflicting sources attempt to influence a target, the source that has greater social impact will be more successful in gaining the target's compliance (Latané & Nida, 1980). The social influence literature provides mixed findings regarding the simultaneous influence of different sources (i.e., the supervisor and subordinate group). For example, in experimental settings, where the authority figure is more salient than a group of individuals, the authority figure has more influence than the group of individuals (e.g., Bocchiaro, Zimbardo, & Van Lange, 2011). In other contexts (i.e., when

dealing with safety behaviors of youth workers), the peer group has more influence than the supervisor as the group's behaviors are more salient, than the perceived preferences of the supervisor (e.g., Westaby & Lowe, 2004). Consistent with social impact theory, this happens when the group has greater immediacy or is in closer proximity and has much more direct contact with the individual.

There is also research that suggests that when conflicting influences are encountered, the individual will consider both sources in the evaluation (e.g., Sobieszek & Webster, 1973). This is consistent with performance appraisal research (e.g., Kozlowski et al., 1998) that suggests raters consider the various constituents involved in the performance appraisal process. Other performance appraisal research (e.g., Mero et al., 2007) also suggests that raters are affected by both the supervisor and the group. Mero and colleagues (2007) conducted an experiment with MBA students who had to rate the performance of undergraduate students. They found that raters provide more accurate ratings when they are accountable (i.e., had to justify their ratings) to those higher in status (i.e., PhD students), than those lower in status (i.e., undergraduate students), compared to themselves. When raters were accountable to ratees, ratings were inflated. It is important to note that participants were not aware of the preferences of either source, but they indicated that they were concerned about both audiences when they were accountable to both.

To investigate whether the rater's supervisor or subordinate group has a greater impact on ratings when inconsistent information is provided, social impact theory is used as it explains when social influence from sources, more and less powerful than the target, has the strongest effects (Latané, 1981). As previously noted, social impact theory proposes that social impact is a multiplicative function of immediacy, strength, and number of the source. Immediacy is likely to

be the same for the rater's supervisor and group, as previously suggested. Given the supervisor's multiple bases of power, ability to administer rewards, and satisfy several rater needs, the supervisor will have greater power (i.e., strength) than the group. The group will have greater number, however, as by definition, the group is a collection of two or more individuals. In this context, strength or power of the source should be more important than the number of the source. This is because strength is especially important when the target is concerned with approval from the source of the influence (Williams & Williams, 1989). This should be the case in this context, as the information presented to the rater's makes approval salient.

Integrating social impact theory with expectancy theory would suggest that in this situation, the supervisor would have greater impact. This is because the supervisor's approval provides more valued outcomes, overall, than the subordinate group's approval, given the supervisor's power relative to the rater, and the group. Additionally, raters in this context should be more concerned with approval than with accuracy, making source strength (i.e., power) especially important. In expectancy theory terms, the valence associated with supervisor approval should be higher than the valence associated with subordinate group approval. Therefore, $\sum[P \rightarrow O(V)]$ for rating consistently with the supervisor's preference will be greater than $\sum[P \rightarrow O(V)]$ for rating consistently with information from the subordinate group. Taken together, the supervisor's information should have a stronger effect on ratings than the subordinate group's information.

Therefore, performance ratings will be higher when the supervisor provides information encouraging high performance ratings and the subordinate group provides information encouraging low performance ratings than when the subordinate group provides information encouraging high performance ratings and the supervisor provides information encouraging low

performance ratings. At the same time, performance ratings will be lower when the rater's supervisor conveys information encouraging low performance ratings and the subordinate group provides information encouraging high performance ratings than when the subordinate group provides information encouraging low performance ratings and the supervisor provides information encouraging high performance ratings. Additionally, the effect of information on performance ratings should be stronger when the supervisor conveys information encouraging inflation and the subordinate group provides information encouraging rating deflation than when information from the supervisor supports deflation and information from the subordinate group encourages deflation. This is because supervisor approval will be valued more than subordinate group approval and more overall positive outcomes will be associated with high performance ratings compared to low performance ratings.

Hypothesis 10a: Performance ratings will be higher when the rater's supervisor provides inflation information and the subordinate group provides deflation information than when the subordinate group provides inflation information and the supervisor provides deflation information.

Hypothesis 10b: Performance ratings will be lower when the rater's supervisor provides deflation information and the subordinate group provides inflation information than when the subordinate group provides deflation information and the supervisor provides inflation information.

Hypothesis 10c: Supervisor inflation information and subordinate group deflation information will have a stronger effect on performance ratings than supervisor deflation information and subordinate group inflation information.

When the rater's supervisor and subordinate group provide inconsistent rating information, the effect on distortion may be weaker or stronger than when only one source provides rating information or when no rating information is provided. We lack a clear understanding of these complex interpersonal relationships affecting the performance appraisal

process, however (Levy & Williams, 2004). Therefore, these interactive effects are examined in the final two sets of hypotheses.

According to social impact theory, the impact of the supervisor on performance ratings should be attenuated by the subordinate group, and vice versa, due to the breaking of consensus and reduction in number of the source. This implies that the effect of the supervisor and the subordinate group will be weaker when conflicting information is encountered than when only one source provides information. For example, the subordinate group should have an effect on performance ratings to a greater extent when the supervisor provides no information regarding performance ratings than when the supervisor prefers ratings in the opposite direction. When the supervisor provides information encouraging rating behavior that is contrary to the subordinate group's information, the supervisor will disapprove of the ratings that are consistent with the information provided by the subordinate group. Therefore, $\sum[P_{\text{inflate}} \rightarrow O(V)]$, for example, will be higher when the subordinate group provides information encouraging high performance ratings and the supervisor provides no information than when the supervisor provides information encouraging low performance ratings. This is because supervisor approval of the ratings will be taken into consideration when the supervisor provides information conveying a rating preference. Therefore, I make the following predictions:

Hypothesis 11a: Performance ratings will be higher when only the supervisor provides inflation information than when the supervisor provides inflation information and the subordinate group provides deflation information.

Hypothesis 11b: Performance ratings will be higher when only the subordinate group provides inflation information than when the subordinate group provides inflation information and the supervisor provides deflation information.

Hypothesis 11c: Performance ratings will be lower when only the supervisor provides deflation information than when the supervisor provides deflation information and the subordinate group provides inflation information.

Hypothesis 11d: Performance ratings will be lower when only the subordinate group provides deflation information than when the subordinate group provides deflation information and the supervisor provides inflation information.

Social impact theory would also suggest that the supervisor's information still has an effect on rater motivation and performance ratings even when conflicting information from the subordinate group is present. Specifically, social impact theory posits that individuals comply with the information from the source that has greater impact. Given the outcomes associated with supervisor approval compared to subordinate group approval, the valence of supervisor approval will be higher than the valence of subordinate group approval. Therefore, the supervisor should have greater impact than the group. For example, when the supervisor provides information encouraging high performance ratings and the subordinate group provides information encouraging low performance ratings, the supervisor will have a stronger effect on ratings than the subordinate group. In such a situation, where the supervisor provides information encouraging high performance ratings and the subordinate group provides information encouraging low performance ratings, ratings should be higher than when no information is provided by either source. The rater's $\sum[P_{\text{inflation}} \rightarrow O(V)]$ will be higher when the supervisor provides information encouraging high performance ratings than when neither source provides information. This is because the rater should experience some positive outcomes (i.e., supervisor approval of ratings) as a result of high performance ratings.

There are no salient positive outcomes when no information is provided by either source, however.

Hypothesis 12a: Performance ratings will be higher when the supervisor provides inflation information and the subordinate group provides deflation information than when no information is provided.

Hypothesis 12b: Performance ratings will be lower when the supervisor provides deflation information and the subordinate group provides inflation information than when no information is provided.

CHAPTER 3

METHODOLOGY

Subjects

Two hundred seventy undergraduate students enrolled in business courses at a southern university in the United States participated in the study for course extra credit. These participants represent a subsample from the larger study ($N = 1123$), and were selected based on the conditions to which they were randomly assigned. 55.9% of the participants were male, 79.3% were white, and the average age was 20.6 years ($SD = 2.4$ years). 47.0% of participants were employed, with an average tenure of 18.9 months ($SD = 23.4$ months), working 19.9 hours per week ($SD = 9.3$ hours). Of all the participants, 45.6% were familiar with performance appraisals. Of those, 63.9% had their performance formally evaluated at work, 24.6% formally rated the performance of a subordinate, 32.0% formally evaluated the performance of a co-worker, and 26.2% formally rated the performance of a supervisor.

Task and Materials

Each participant was asked to assume the role of a manager at a local bank and rate the job performance of Julie, one of his or her bank tellers (see Appendix B). Prior to rating the bank teller's performance, participants watched a training video (22 minutes and 39 seconds), intended to provide the prerequisite knowledge required to evaluate the teller's performance. In the training video, the three performance dimensions (i.e., money handling, cross-selling, and professional conduct) the bank teller was supposed to be rated on were explained. Additionally, examples of above average, average, and below average performance for each performance

dimension were depicted. After completing the training, participants watched and rated three performance scenarios (see Appendix C), each lasting between 24 and 47 seconds, for practice.

For the actual rating task, participants watched six videotaped performance scenarios involving an interaction between the bank teller and a bank customer and evaluated the teller's performance. Performance scenarios captured at least one of the three performance dimensions (i.e., money handling, cross-selling, and professional conduct) and lasted between 59 and 80 seconds. The scripts for each of the six rating scenarios, along with total duration, are included in Appendix D.

Development of Rating Scales

Participants used a graphic rating scale (GRS), ranging from 1 (low) to 7 (high) to evaluate the bank teller's performance. The scale also included a "not applicable" option in case a specific performance dimension was not captured in the performance scenario. The scale was developed to complement the performance dimensions used as criteria against which the bank teller's performance was evaluated. The performance dimensions were defined as follows:

Money Handling: Handling money properly entails counting money out loud, quickly and accurately. It also involves sorting bills into stacks according to their denominations, placing each bill into the proper stack, and making sure that each bill faces up and in the same direction.

Cross-Selling: With cross-selling, tellers pay attention to the needs of their individual customers and—when appropriate—suggest a product or service to the customer that would help better serve those needs.

Professional Conduct: A bank teller has to behave professionally with customers at all times. A teller acts in a professional manner by being business-like and friendly, using appropriate language, and making the customer feel welcome.

Several items capturing each performance dimension in detail, an item assessing overall performance for each dimension, and an item measuring overall performance for the scenario were used. All items are listed in Appendix E.

Development and Taping of Performance Scenarios

The performance scenarios were written such that each performance dimension, for each scenario, had an intended true score, ranging from 1 to 7. Additionally, each performance scenario was written to depict overall average performance (i.e., 4). During the writing of the scripts, two management faculty and doctoral students discussed all performance dimensions and agreed on the true scores for dimensions and scenarios. Actors were used to play all the roles. One actor played the role of the bank teller that that needed to be evaluated. Six different actors, one for each performance scenario, played the role of bank customers. Actors closely followed the script provided in Appendix D. The true score estimates were discussed again, and agreed upon, after the performance scenarios were watched on video by the two faculty and PhD students.

Validation of True Score Estimates

To validate the true score estimates, expert raters were used. Seven individuals familiar with the performance appraisal process and performance rating (i.e., five OB/HR PhDs and two OB/HR doctoral students) served as expert raters. All expert raters completed a frame-of-reference training (Bernardin & Buckley, 1981). Specifically, all expert raters watched a training

video that explained the criteria against which a bank teller's job performance is evaluated as well as the definitions and components of each dimension. As part of the training video, raters saw examples of high, medium, and low performance on each dimension. Additionally, expert raters were able to ask questions and engage in discussion with the trainers in a group setting. This ensured that all raters had a common reference standard for evaluating performance (Woehr, 1994).

After this initial discussion, all expert raters independently rated three practice scenarios. Consistent with Bernardin and Buckley's (1981) recommendations, appropriate performance ratings for each of the practice scenarios were discussed by the trainers. This further ensured a common frame of reference for the expert raters. Finally, raters watched and independently rated the six performance scenarios. Raters had the option to take notes. Expert raters used both the GRS and behaviorally-anchored rating scale (BARS). The BARS was specifically developed for this study and is included in Appendix F. The GRS was similar to the one used by the participants in the study but included additional items. The GRS the experts used is included in Appendix G.

Dependent Variables

Cronbach (1995) proposed that accuracy is a multidimensional construct, consisting of four components that capture different aspects of accuracy and provide different types of information about ratings. The elevation (E) component represents the difference between the mean rating and the mean true score, averaged across all ratees and all items (Sulsky & Balzer, 1988). Elevation is useful in determining whether the rater is being lenient or harsh compared to true score estimates (Kline & Sulsky, 2009). Essentially, E captures the extent to which

performance ratings are distorted. The direction of distortion, however, is not taken into consideration by the elevation component (see formula for calculating Elevation in Appendix H). Leniency/severity (McIntyre et al., 1984) retains information about directionality, and while it is more descriptive than E, when squared, it produces calculations identical to Cronbach's E^2 (Sulsky & Balzer, 1988).

Research (e.g., Murphy, Garcia, Kerkar, Martin, & Balzer, 1982; Stamoulis & Hauenstein, 1993) focusing on rating accuracy has used these operationalizations of accuracy, as there was variance across performance true score estimates across different ratees. Stated differently, true score estimates of performance were not identical across ratees. The present study focuses on the performance of one ratee across a number of performance dimensions, and the true score for overall performance for each scenario reflects average performance. As the true score reflects a constant, the dependent variable used in this study is the overall performance rating, averaged across the performance scenarios. This operationalization provides a number of advantages. One, the direction of distortion can be determined as the average rating can be compared to the true score estimate (i.e., 4). Two, this does not require any transformations, therefore the dependent variable remains in original units. Three, it is consistent with recent performance appraisal research focusing on rating accuracy (e.g., Judge & Ferris, 1993; Spence & Keeping, 2010).

Procedures

Participants were recruited in undergraduate business courses. During recruitment, participants were informed that they will earn course extra credit for completing the study and that they would have an opportunity to earn between \$0 and \$20 by participating in the study. Participants signed up for the study through Sona systems (an online system used to coordinate

participation in research). Once in the laboratory, participants were greeted, asked to complete the informed consent form, and signed in. Random assigned was used to assign participants into experimental conditions. Participants were instructed to go to the computer room, go to the assigned laptop number, and begin the study. Another experimenter was in the computer room and directed participants to their designated laptop stations. Once at the appropriate laptop station, each participant put on headphones and began the study by watching an introductory video (2 minutes and 24 seconds). During the introductory video, participants were introduced to the study, and informed that they would be rating the performance of a bank teller, watching a training video, and completing a number of questionnaires.

After watching the introduction video, each participant entered his or her participant identification number and was directed to an overview of the study. In the overview, they were given an outline of the study schedule (see Appendix I) to have a better understanding of what they would be doing throughout the study. Each participant then completed a questionnaire, followed by the description of their role in the study, the training video, and another questionnaire. Participants then had the opportunity to practice rating scenes. Specifically, each participant watched a practice rating scenario, was given the opportunity to rate the teller's performance, and then watched another video explaining the performance ratings that were merited in that particular performance scenario. Then, participants completed a third questionnaire prior to the experimental manipulations.

Participants then read the experimental manipulations. After the manipulations, they completed a fourth questionnaire and then watched and provided ratings for the six performance scenarios. Specifically, participants watched each rating scenario and provided performance ratings for that scenario immediately after watching it. This was done for each of the six rating

scenarios. Participants responded to the items illustrated in Appendix E. To control for possible order effects, the rating scenarios were presented in six random orders. Finally, participants completed one final questionnaire and read a debriefing. The entire study was completed at the laptop stations. Participants went to another room in the laboratory, were debriefed, given an opportunity to ask questions, and answered some final follow-up questions. Participants spent a total of 60 to 90 minutes in the laboratory.

Experimental Manipulations

The present study is part of a larger study that was a 3 (supervisor information) x 3 (subordinate group information) x 4 (incentive pay) between-subjects design with thirty-six experimental conditions. The supervisor information manipulation had three levels. The supervisor either provided information expressing no preference regarding the bank teller's performance ratings, information expressing a preference for high performance ratings, or information expressing a preference for low performance ratings. The subordinate group manipulation also had three levels. The subordinate group either provided no information regarding the bank teller or expressed an opinion that the bank teller is a good employee or a bad employee. The incentive pay manipulation had four levels: no incentive pay, low incentive pay, medium incentive pay, and high incentive pay. The present study focused on supervisor and group information and therefore utilized the nine no incentive pay conditions. The text for the supervisor and group manipulations is included in Appendix J.

Manipulation Checks

Manipulation checks were conducted as part of a pilot study and during the current study to examine the effectiveness of the manipulations. A total of 12 items was used to assess the

effectiveness of the supervisor and subordinate group manipulations. For both the current study and the pilot study, participants responded to all 12 items on a 1 (Strongly Disagree) to 5 (Strongly Agree) scale.

6 items focused on the supervisor manipulation. Of these 6 items, 3 items measured the extent to which participants perceived that the supervisor wants them to provide high performance ratings for the bank teller ($\alpha = .83$ for the current study, $\alpha = .74$ for the pilot study). Items included: “Pat (my boss) wants me to give Julie a good rating,” “I will please my boss if I rate Julie high,” and “To make my boss happy, I need to give Julie a high rating.” The other three items (i.e., “I will satisfy my boss if I give Julie a low rating,” “Pat will like it if Julie gets a bad rating,” “To satisfy Pat, I need to give Julie a low rating”) were used to assess the extent to which participants perceived that the supervisor prefers low performance ratings for the bank teller ($\alpha = .85$ for the current study, $\alpha = .84$ for the pilot study).

The other 6 items were used to evaluate the subordinate group manipulation. Three items measure the extent to which participants perceived that the subordinate group felt the bank teller, Julie, is a good employee ($\alpha = .96$ for the current study, $\alpha = .94$ for the pilot study). Items included: “Her coworkers think Julie is a good worker,” “Julie is considered a good performer by her team,” and “The other tellers describe Julie as being productive.” The remaining three items (i.e., “Julie’s coworkers describe her as being unproductive,” “Her coworkers think of Julie as a bad employee,” “The other tellers think that Julie does a bad job”) assess the extent to which participants perceived that the group thinks Julie is a bad employee ($\alpha = .96$ for the current study, $\alpha = .84$ for the pilot study).

Pilot Study Manipulation Checks

To assess the effectiveness of the supervisor and subordinate group manipulations for the pilot study, four One-Way ANOVAs were conducted. Two One-Way ANOVAs were conducted with the supervisor manipulation (i.e., high ratings, low ratings) as the between-subjects factor. The results indicate that the supervisor manipulation had an effect on participants' perceptions of the supervisor wanting them to provide high performance ratings for the bank teller, $F(1,61) = 52.36, p < .001$. Results of the second One-Way ANOVA indicate that the supervisor manipulation also had an effect on participants' perceptions of the supervisor wanting them to provide low performance ratings, $F(1,61) = 72.91, p < .001$, for the bank teller. The results of both ANOVAs focusing on the supervisor manipulation are displayed in Table 1 below. The means for the manipulation checks across the supervisor manipulation conditions are displayed in Table 2 on the following page.

Table 1

Effects of Supervisor Information Manipulation on Participants' Perceptions of Supervisor Rating Preferences

Measure	Source	df	MS	F
Preference for High Ratings	Manipulation	1	28.12	52.36**
	Error	61	0.54	
Preference for Low Ratings	Manipulation	1	40.49	72.91**
	Error	61	0.56	

Note: ** $p < .01$

Table 2

Means for Manipulation Checks Across Supervisor Information Manipulations

Manipulation Check	Supervisor Information	
	Low Ratings	High Ratings
Preference for High Ratings	2.09	3.43
Preference for Low Ratings	3.79	2.19

Note: $N = 63$.

Two additional ANOVAs were conducted with the subordinate group manipulation (i.e., good ratee, bad ratee) as the between-subjects factor. The results indicate that the subordinate group manipulation had an effect on participants' perceptions of the subordinate group seeing Julie as a good employee, $F(1,61) = 318.84$, $p < .001$, and a bad employee $F(1, 61) = 92.24$, $p < .001$. The results of the two ANOVAs are displayed in Table 3 on the following page. Means for manipulation checks for subordinate group information are contained in Table 4 on the following page. The results of the pilot study indicate the manipulations were effective overall.

Table 3

Effects of Subordinate Group Information Manipulation on Participants' Perceptions of the Subordinate Group's Opinion of the Ratee

Measure	Source	df	MS	F
Good Employee	Manipulation	1	101.84	318.84**
	Error	61	0.32	
Bad Employee	Manipulation	1	59.05	92.24**
	Error	61	0.64	

Note: ** $p < .01$

Table 4

Means for Manipulation Checks for Subordinate Group Information Manipulations

Manipulation Check	Subordinate Group Information	
	Low Ratings	High Ratings
Preference for High Ratings	2.01	4.55
Preference for Low Ratings	3.64	1.70

Note: $N = 63$

Manipulation Checks for Current Study

As previously noted, manipulation checks were also conducted during the present study. An ANOVA was conducted with the supervisor manipulation (i.e., no rating information, high ratings, low ratings) as the between-subjects factor. The results indicate that the supervisor

manipulation had an effect on participants' perceptions of the supervisor wanting them to provide high performance ratings, $F(2, 267) = 40.21, p < .001$, and low performance ratings, $F(2, 267) = 79.02, p < .001$, for the bank teller. The results of the ANOVA are displayed in Table 5 below.

Table 5

Effects of Supervisor Information Manipulation on Participants' Perceptions of Supervisor Rating Preferences

Measure	Source	df	MS	F
Preference for High Ratings	Manipulation	2	23.78	40.21**
	Error	269	0.59	
Preference for Low Ratings	Manipulation	2	46.91	79.02**
	Error	269	0.59	

Note: ** $p < .01$

Tukey's test was conducted to examine which group means were significantly different. Results show that participants perceived the supervisor wanted them to provide high performance ratings for the bank teller to a greater extent in the high ratings condition ($M = 3.05, SD = .89$) than in the low ratings condition ($M = 2.01, SD = 0.64$) or the no rating information condition ($M = 2.47, SD = 0.76$). Participants also perceived the supervisor will be more pleased with high performance ratings in the no rating information condition than in the low ratings

condition. Furthermore, participants perceived the supervisor would be satisfied with low performance ratings to a greater extent in the low ratings condition ($M = 3.39$, $SD = .86$) than the high ratings condition ($M = 2.11$, $SD = 0.70$) or the no rating information condition ($M = 2.16$, $SD = 0.74$). The means for the manipulation checks across supervisor information manipulations are displayed in Table 6 below.

Table 6

Means for Manipulation Checks Across Supervisor Information

Manipulation Check	Supervisor Information		
	None	Low Ratings	High Ratings
Preference for High Ratings	2.47	2.01	3.04
Preference for Low Ratings	2.16	3.39	2.11

Another ANOVA was conducted with the subordinate group manipulation (i.e., no ratee information, good ratee, bad ratee) as the between-subjects factor. The results indicate that the subordinate group manipulation had an effect on participants' perceptions of the subordinate group seeing Julie as a good employee, $F(2, 267) = 344.44$, $p < .001$, and a bad employee $F(1, 267) = 354.68$, $p < .001$. The results of the multivariate ANOVA are displayed in Table 7 on the following page.

Table 7

Effects of Subordinate Group Information Manipulation on Participants' Perceptions of the Subordinate Group's Opinion of the Ratee

Measure	Source	df	MS	F
Good Employee	Manipulation	1	101.84	318.84**
	Error	61	0.32	
Bad Employee	Manipulation	1	59.05	92.24**
	Error	61	0.64	

Note: ** $p < .01$

Tukey's test indicates that participants perceived that the subordinate group viewed the bank teller as a good employee to a greater extent in the good ratee condition ($M = 4.44$, $SD = .70$) than the bad ratee condition ($M = 1.95$, $SD = 0.68$) or the no ratee information condition ($M = 3.31$, $SD = 0.51$). Participants also perceived that the subordinate group sees Julie as a good employee to a greater extent in the no ratee information condition than the bad ratee condition. Participants perceived that the group views the bank teller as a poor employee to a greater extent in the bad ratee condition ($M = 3.88$, $SD = .67$) than the good ratee condition ($M = 1.44$, $SD = 0.63$) and the no ratee information condition ($M = 2.64$, $SD = 0.52$). Finally, participants also perceived that the subordinate group sees Julie as a bad employee to a greater extent in the no ratee information condition than the good ratee condition. Table 8 (illustrated on the next page) displays the mean responses to the manipulation checks across subordinate group manipulations.

Table 8. Means for Manipulation Checks Across Subordinate Group Information Manipulations

Manipulation Check	Subordinate Group Information		
	None	Bad Employee	Good Employee
Good Employee	3.31	1.94	4.44
Bad Employee	2.64	3.88	1.44

CHAPTER 4

RESULTS

Evaluation of Interrater Agreement

To ensure agreement regarding the true performance for each performance scenario, interrater agreement for the expert raters was calculated using the average deviation (AD) index (Burke, Finkelstein, & Dusig, 1999). The AD index assesses response dispersion across the mean or median response, such that smaller values indicate greater agreement (Burke & Dunlap, 2002). AGREE, a program available at <http://www.tulane.edu/~dunlap/psylib.html> was used to conduct the analyses (Burke & Dunlap, 2002; Dunlap, Burke, & Smith-Crowe, 2003). While r_{WG} -type interrater agreement indices have been used most commonly in previous work, research (e.g., Dunlap et al., 2003) suggests that these have some limitations (e.g., specification of the null distribution) and that the AD index offers a more appropriate alternative. Researchers (e.g., Burke & Dunlap, 2002), however, have suggested that several interrater agreement indices should be used to make the most informed decision regarding rater agreement. As such, both AD index for the mean and median as well as r_{WG} were calculated using AGREE.

An average deviation (AD) equal to zero would suggest perfect agreement (Burke & Dunlap, 2002). Burke and Dunlap (2002) identified $c/6$ (where c is the number of response options) as the upper-limit cut-off for the AD index. Specifically, if the AD mean or median value is below $c/6$, there is sufficient evidence to conclude that there is interrater agreement. With seven response categories, values below 1.167 (i.e., $7/6$) are statistically significant at $p < .05$ and indicative of interrater agreement. Dunlap et al. (2003) also take the number of raters into consideration, such that the average deviation needs to be lower to indicate significant

interrater agreement when there are fewer raters. In other words, AD values need to be closer to 0 to indicate agreement as the number of raters decreases. When the number of raters is equal to seven, AD values equal to or less than .86 are both statistically and practically significant at $p < .05$ (Dunlap et al., 2003).

When multiple ratees are evaluated across a number of performance dimensions, the ratings for all performance dimensions across all ratees would be aggregated to arrive at leniency/severity and the elevation component of accuracy. In the present study, raters only evaluate one ratee across six performance dimensions. Additionally, the performance true score estimates for the performance dimensions range from 1 to 7. Performance scenarios contained at least one performance dimension (i.e., professional conduct), but could contain up to three performance dimensions⁵. If the performance true score estimate for a specific dimension is 1, this does not give raters the opportunity to deflate performance ratings. Alternatively, if the performance true score estimate for a performance dimension is 7, raters do not have the opportunity to inflate performance ratings. Each performance scenario, however, should have a true score estimate equal to 4, depicting average performance. This provides raters with the opportunity to both inflate and deflate performance ratings. Given these conditions, the overall rating for each scenario is used instead of the individual performance dimensions.

To evaluate true score estimates, interrater agreement was calculated using the AD and rwG-type interrater agreement indices for each performance scenario. Indices were calculated based on the GRS ratings experts provided. The results of these analyses are presented in Table 9 on the following page. Both indices provide evidence for acceptable levels of interrater

⁵ In addition to the professional conduct dimension, Scene 1, 4, and 6 also captured the money handling dimension. Scenes 2 and 5 captured the professional conduct and cross-selling dimensions. Scene 3 only captured the professional conduct dimension.

agreement for Scenes 1, 2, 3, 4, and 5. Interrater agreement for Scene 6 did not reach an acceptable level according to both indices. Specifically, the r_{WG} index was below the generally accepted .70 cut-off for agreement (Lance, Butts, & Michaels, 2006). And the AD values were not practically significant as they were greater than .86. Therefore, Scene 6 will not be used to test the hypotheses.

Table 9. Interrater Agreement for Performance Scenarios

	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5	Scene 6
True Score Estimate Information						
Median Rating	2.00	4.00	4.00	4.00	4.00	4.00
Mean Rating	2.43	3.43	4.00	3.86	3.86	4.14
SD	0.98	0.79	0.82	0.69	0.90	1.35
r_{WG}	0.76*	0.85**	0.83**	.88**	0.80*	0.55†
$AD_{(median)}$	0.71*	0.57**	0.57**	.43**	.71*	1.00†
$AD_{(mean)}$	0.78*	0.65**	0.57**	.49**	.74*	1.02†

Note: ** $p < .01$

* $p < .05$

† $p < .10$

The median expert rating for Scenes 2, 3, 4, and 5 was equal to 4. This is consistent with the intended true score for each of these scenarios. The median rating for Scene 1 was 2. This is considerably lower than the intended true score estimates and provides less opportunity for participants to deflate performance ratings for this scenario. Therefore, Scenes 2, 3, 4, and 5, which had overall performance true score estimates equal to 4, were used to calculate the overall performance rating.

Tests of Hypotheses

Prior to testing the proposed hypotheses, the data were examined for outliers, using recommendations for identifying and handling potential outliers (Aguinis, Gottfredson, & Joo, 2013). The purpose of this was to determine if any participants provided overall mean performance ratings that deviated from the mean for their experimental group. As recommended, the data were first examined visually. Specifically, visual aids including Q-Q plots, Stem-and-Leaf plots, and box plots were used. Then, the data were numerically examined. Specifically, the standard deviation of each participant's mean rating was calculated based on the mean rating and standard deviation of the group the participant was randomly assigned to. As recommended by Aguinis and colleagues (2013), responses ± 2.24 standard deviations from the mean were removed. This resulted in the removal of eight participants' mean ratings from six of the nine experimental conditions⁶.

3 x 3 Factorial ANOVA

A 3 (no supervisor information, supervisor inflation information, supervisor deflation information) x 3 (no subordinate group information, subordinate group inflation information, subordinate group deflation information) factorial ANOVA was conducted to examine the main effects of supervisor and group information and the interaction effect. Results indicate that supervisor information did not have an effect on performance ratings, $F(2, 253) = 1.245, ns$.

⁶ One response was removed from each of the following conditions: no subordinate group information and supervisor inflation information, no supervisor information and subordinate group deflation information, no supervisor information and subordinate group inflation information, supervisor and subordinate group deflation information, supervisor deflate and subordinate group inflate. Three responses (two that fell 2.24 standard deviations below and one that fell 2.24 standard deviations above the group mean) were removed from the supervisor and subordinate group inflate condition.

Subordinate group information also did not have an effect on performance ratings, $F(2, 253) = 2.175$, *ns*. Finally, the interaction between supervisor and subordinate group information was not significant, $F(4, 253) = 1.771$, *ns*. Table 10, below, contains all cell means, marginal means, and the grand mean.

Table 10

Cell Means, Marginal Means, and Grand Means

Subordinate Group Manipulation	Supervisor Manipulation			Total
	None	Low	High	
None	3.74	3.98	3.90	3.87
Low	3.70	3.77	3.88	3.78
High	4.13	3.72	4.20	4.01
Total	3.86	3.82	3.99	3.89

The hypotheses in the present study focus on comparisons between specific cells and differences in effect size magnitudes. They are also planned comparisons, hypothesized apriori. Given these conditions, it is not necessary that the overall main effects and interaction effect are significant. To test the specific proposed hypotheses, additional analyses are required.

Specifically, predictions regarding cell means were examined using One-way ANOVAs and post-hoc tests.⁷ To test differences in effect size magnitudes, χ^2 tests were used.

Inflation versus Deflation Information

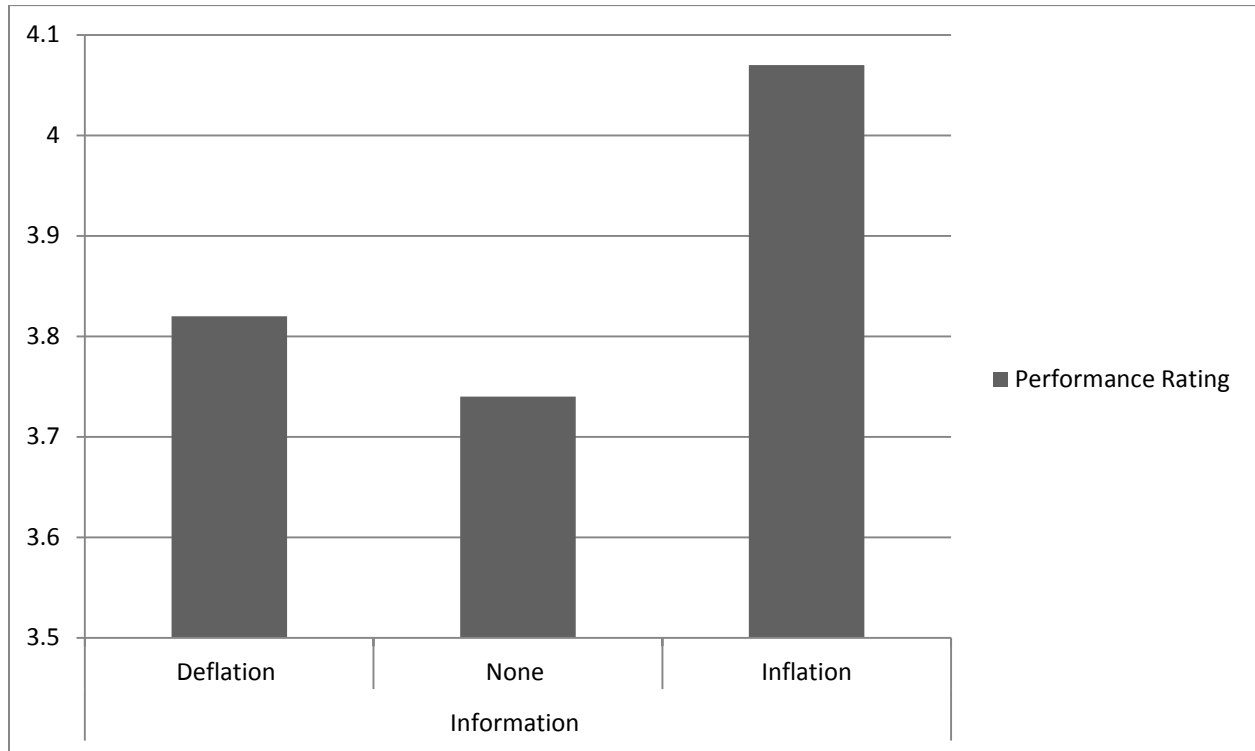
Hypothesis 1 predicted that inflation information will have a stronger effect on performance ratings than deflation information. To test this hypothesis, a significance test comparing the effect sizes of inflation and deflation information was performed. Specifically, an inferential test developed by Fowler (1987) was used to compare the magnitude of effects of inflation and deflation information on performance ratings. The formulas used to conduct the significance test are illustrated in Appendix K. To obtain effect sizes, two One-way ANOVAs were conducted, one with inflation information (inflation information, no information) and one with deflation information (deflation information, no information). The results indicate that performance ratings in the inflation information conditions ($M = 4.07$, $SD = 0.75$) were significantly greater, $F(1, 113) = 4.147$, $p < .05$, $\eta^2 = .035$, than ratings in the no information condition ($M = 3.74$, $SD = .83$). Performance ratings in the deflation information condition ($M = 3.82$, $SD = 0.75$) were not significantly different, $F(1, 116) = .218$, ns , $\eta^2 = .002$, from ratings in the no information condition ($M = 3.74$, $SD = .83$). The χ^2 test indicates that the effect magnitude of inflation information on performance ratings was significantly different from the effect magnitude of deflation information on performance ratings, $\chi^2(1) = 2.84$, $p < .05$ (one-tailed test). Therefore, Hypothesis 1 was supported. A bar chart displaying these effects is illustrated in Figure 1 on the following page.

⁷ Tukey's test was used to compare differences between cell means. According to Tukey's test, the difference between means has to be greater than .388 for the means to be significantly different. The cut-off value is based on 9 experimental conditions, an average of 29.11 participants per cell, a mean square error of 0.560, and 253 degrees of freedom.

Figure 1

Effects of Deflation Information Versus Inflation Information on Performance Ratings

(Hypothesis 1)



Main Effects of Supervisor

Hypothesis 2a predicted that performance ratings will be lower when the supervisor provides information encouraging low performance ratings than when no information is provided. To test Hypothesis 2a, a One-way ANOVA was conducted with supervisor information (none, low ratings) as the between-subjects factor. Results indicate that ratings were not significantly different [$F(1, 58) = 1.230, ns$] in the low ratings condition ($M = 3.98, SD = .86$) than the no rating information condition ($M = 3.74, SD = 0.83$). Additionally, Tukey's

follow-up test indicates that the means were not significantly different, as the mean difference is 0.24, which is below the 0.388 cut-off. Therefore, Hypothesis 2a was not supported.

For Hypothesis 2b, it was posited that performance ratings will be higher when the supervisor provides information encouraging high performance ratings compared to when no information is provided. Hypothesis 2b was also tested using a One-way ANOVA, with supervisor information (none, high ratings) as the between-subjects factor. Results show that the supervisor information did not have a significant effect on performance ratings, $F(1, 57) = 0.572$, *ns*, as ratings were not significantly higher in the high ratings condition ($M = 3.90$, $SD = 0.74$) than in the no rating information condition ($M = 3.74$, $SD = .83$). Tukey's test also indicates that the difference between the means (0.16) was not significantly different. Thus, Hypothesis 2b did not receive support.

Hypothesis 2c proposed that supervisor inflation information will have a stronger effect on performance ratings than supervisor deflation information. The test used for comparing effect magnitudes indicates that the effect magnitude of supervisor inflation information on performance ratings was not significantly different from the effect magnitude of supervisor deflation information on performance ratings, $\chi^2(1) = 2.57$, *ns*. Therefore, Hypothesis 2c was not supported.

Main Effects of Subordinate Group

Hypothesis 3a predicted that performance ratings will be lower when the subordinate group provides information encouraging low performance ratings than when no information is provided. To test Hypothesis 3a, a One-way ANOVA was conducted with subordinate group information (none, bad ratee) as the between-subjects factor. The results indicate that

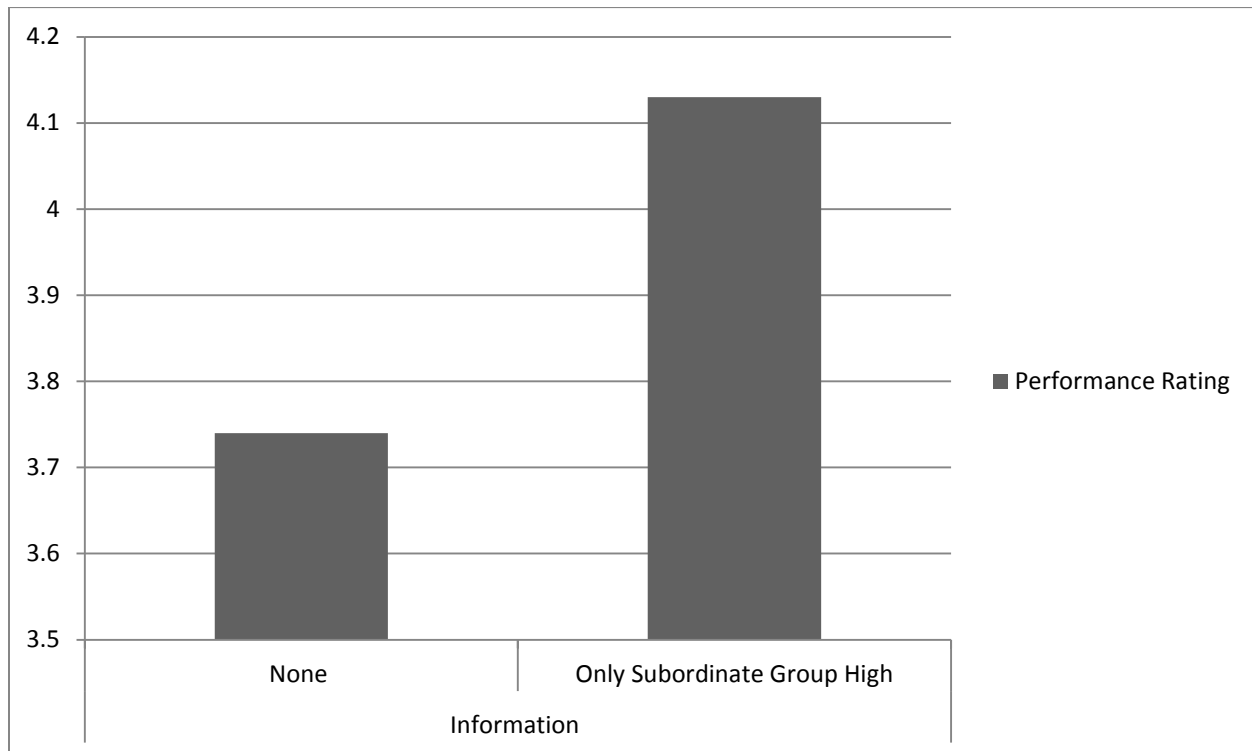
performance ratings in the bad ratee condition ($M = 3.70$, $SD = 0.71$) were not significantly different, $F(1, 57) = .047$, ns , from ratings in the no ratee information condition ($M = 3.74$, $SD = .83$). Additionally, Tukey's test indicates that the difference between the means (0.04) was not significantly different. Therefore, the results do not provide support for Hypothesis 3a.

Hypothesis 3b stated that performance ratings will be higher when the subordinate group provides information encouraging high performance ratings than when no information is provided. To test Hypothesis 3b, another One-way ANOVA, with subordinate group information (none, good ratee) as the between-subjects factor, was conducted. Results indicate that subordinate group information had a significant effect on performance ratings, $F(1, 57) = 3.334$, $p < .05$ (one-tailed test). Performance ratings were significantly higher in the good ratee condition ($M = 4.13$, $SD = .80$) than in the no ratee information condition ($M = 3.74$, $SD = 0.83$). Tukey's test also suggests that the difference between the means (0.39) was significantly different from 0. As such, Hypothesis 3b was supported. A bar chart displaying this effect is illustrated in Figure 2 on the following page.

Figure 2

Effect of Subordinate Group Inflation Information Versus No Information on Performance

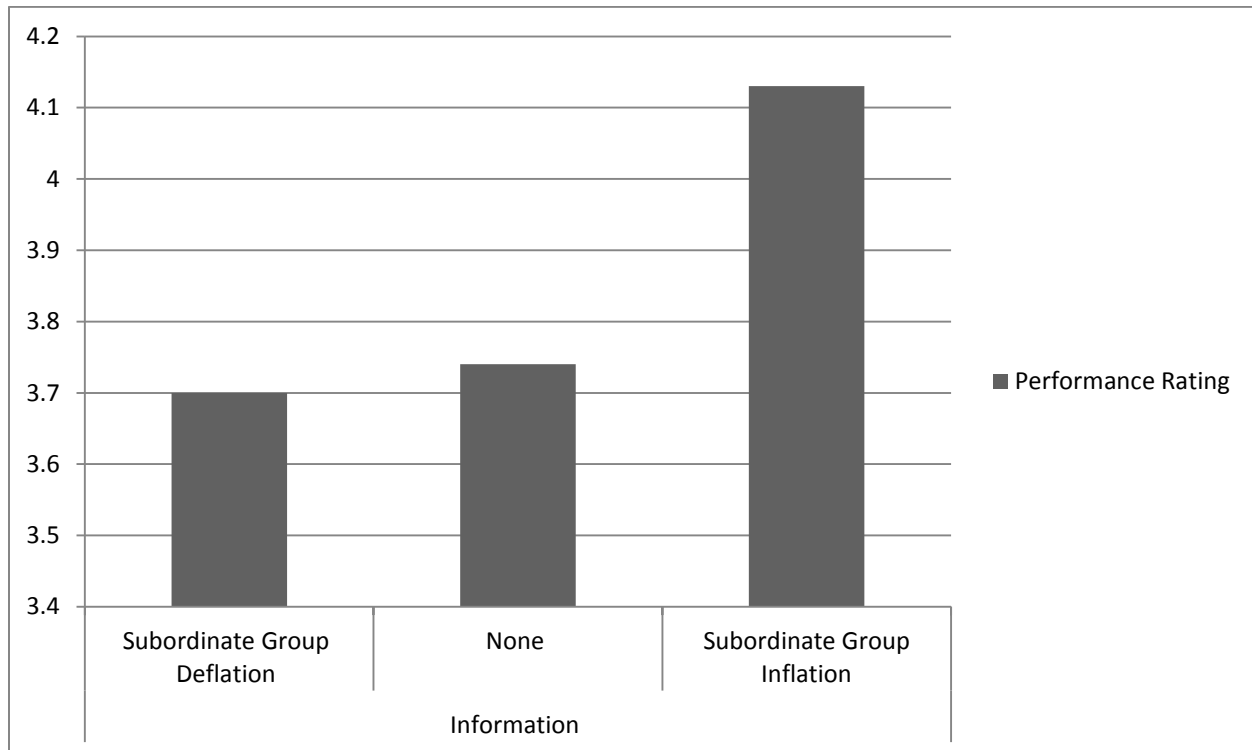
Ratings (Hypothesis 3b)



Hypothesis 3c proposed that subordinate group inflation information will have a stronger effect on performance ratings than subordinate group deflation information. The test used for comparing effect magnitudes indicates that the effect magnitude of subordinate group inflation information on performance ratings was significantly different from the effect magnitude of subordinate group deflation information on performance ratings, $\chi^2(1) = 2.82, p < .05$ (one-tailed test). Therefore, Hypothesis 3c was supported. A bar chart depicting these effects can be found in Figure 3 on the following page.

Figure 3

Effects of Subordinate Group Deflation Information Versus Subordinate Group Inflation Information on Performance Ratings (Hypothesis 3c)



Interaction Effects of Supervisor and Subordinate Group

Hypothesis 4a predicted that performance ratings will be higher when both the supervisor and subordinate group provide information encouraging high performance ratings than when only the supervisor provides information encouraging high performance ratings. To test Hypothesis 4a, a One-way ANOVA was conducted with source of information encouraging inflation (supervisor and subordinate group, supervisor only) as the between-subjects factor. The results indicate that the source of the information did not have a significant effect on performance ratings, $F(1, 54) = 2.619, ns$. While not significant, performance ratings were higher

when both the supervisor and subordinate group provided information encouraging high performance ratings ($M = 4.20$, $SD = .67$) than when only the supervisor preferred high performance ratings ($M = 3.90$, $SD = 0.74$). Tukey's test also indicates that the difference between the cell means (0.30) was not significantly different. Hypothesis 4a was therefore not supported.

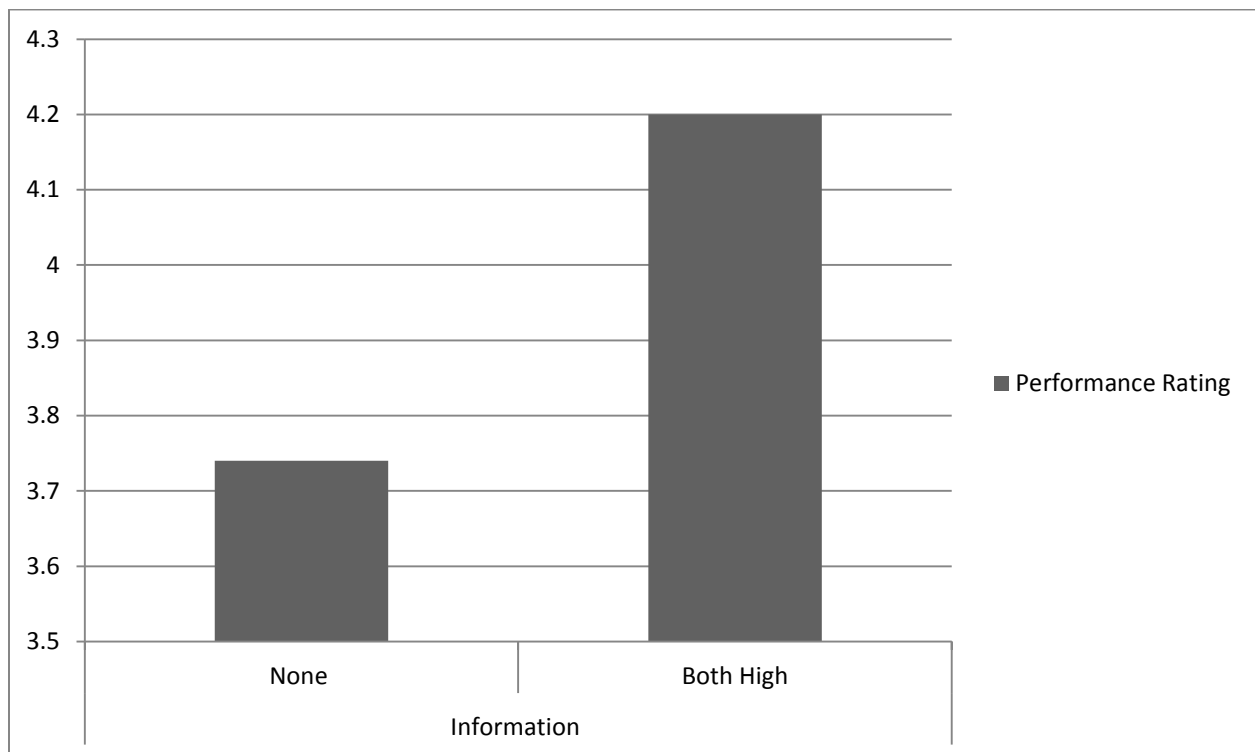
In Hypothesis 4b, it was predicted that performance ratings will be higher when both the supervisor and subordinate group provide consistent information encouraging high performance ratings than when only the subordinate group provides information encouraging high performance ratings. To test Hypothesis 4b, another One-way ANOVA was conducted with source of information prompting high performance ratings (supervisor and subordinate group, subordinate group only) as the between-subjects factor. The results indicate that the source of information did not have a significant effect on distortion, $F(1, 54) = .140$, *ns*, as performance ratings were not significantly higher when both the supervisor and subordinate group provided information prompting high performance ratings ($M = 4.20$, $SD = .67$) than when only the subordinate group provided information encouraging high performance ratings ($M = 4.13$, $SD = .80$). Tukey's test also indicates that the difference between the cell means (0.07) was not significantly different. As such, Hypothesis 4b was not supported.

Hypothesis 4c posited that performance ratings will be higher when both the supervisor and the subordinate group provide information encouraging high performance ratings than when no information is provided. Hypothesis 4c was also tested using a One-way ANOVA, with presence of information encouraging high performance ratings (supervisor and subordinate group, none) as the between-subjects factor. The results indicate that the presence of information had a significant effect on performance ratings, $F(1, 55) = 5.286$, $p < .05$. Performance ratings

were higher when both the supervisor and subordinate group provided information encouraging high performance ratings ($M = 4.20$, $SD = .67$) than when no information was provided ($M = 3.74$, $SD = 0.83$). Additionally, Tukey's test indicates that the difference between cell means (0.46) was significantly different as it exceeds the 0.388 cut-off. As such, Hypothesis 4c was supported. A bar chart displaying this effect is illustrated in Figure 4 below.

Figure 4

Effects of No Information Versus Supervisor and Group Inflation Information on Performance Ratings (Hypothesis 4c)



In Hypothesis 5a, it was predicted that performance ratings will be lower when both the supervisor and the subordinate group provide information encouraging low performance ratings than when only the supervisor provides information encouraging low performance ratings. To test Hypothesis 5a, a One-way ANOVA was conducted with source of information encouraging low performance ratings (supervisor and subordinate group, supervisor only) as the between-subjects factor. The results indicate that the source of information did not have a significant effect on distortion, $F(1, 57) = 1.134, ns$. While not significant, performance ratings were lower when both encouraged low performance ratings ($M = 3.77, SD = .68$) than when only the supervisor provided information encouraging low performance ratings ($M = 3.98, SD = 0.86$). Tukey's test also suggests that the difference between the two means (0.21) was not significantly different. Hypothesis 5a was not supported.

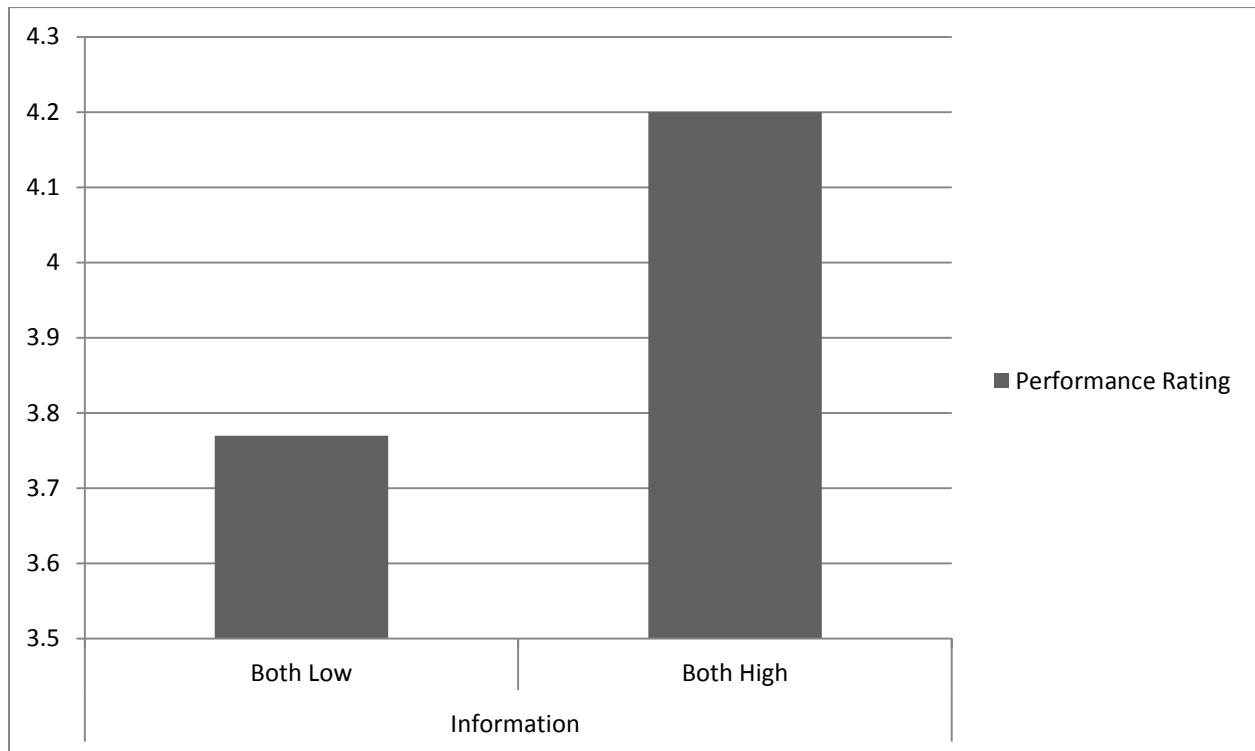
Hypothesis 5b stated that performance ratings will be lower when both the supervisor and subordinate group provide information encouraging low performance ratings than when only the subordinate group provides information encouraging low performance ratings. To test Hypothesis 5b, a One-way ANOVA was conducted with source of information encouraging low performance ratings (supervisor and subordinate group, subordinate group only) as the between-subjects factor. Results show that the source of information did not have a significant effect on performance ratings, $F(1, 56) = .143, ns$, as distortion was not significantly different when the subordinate group thought the ratee is a bad employee ($M = 3.70, SD = .71$) than when both the supervisor and subordinate group provided information prompting low performance ratings ($M = 3.77, SD = 0.68$). Additionally, Tukey's test suggests that the difference between the two means (0.07) was not significantly different. Hypothesis 5b was therefore not supported.

In Hypothesis 5c, it was predicted that performance ratings will be lower when both the supervisor and subordinate group provide information encouraging low performance ratings than when no information is provided. Another One-way ANOVA was conducted to test Hypothesis 5c, with presence of information encouraging low performance ratings (supervisor and subordinate group, none) as the between-subjects factor. The results indicate that the presence of information prompting low performance ratings did not affect performance ratings, $F(1, 57) = .017, ns$, as performance ratings were not significantly lower when both the supervisor and subordinate group provided information encouraging low performance ratings ($M = 3.77, SD = 0.68$) than when no information was provided ($M = 3.74, SD = .83$). Tukey's test also suggests that the difference between the two means (0.03) was not significantly different. Therefore, Hypothesis 5c did not receive support.

In Hypothesis 6, it was predicted that performance ratings will be higher when both the supervisor and subordinate group provide inflation information than information encouraging low performance ratings. A One-way ANOVA was conducted with type of information from both the supervisor and subordinate group (encouraging high performance ratings, encouraging low performance ratings) as the between-subjects factor, to test Hypothesis 6. Results show that the type information from the supervisor and subordinate group had a significant effect on performance ratings, $F(1, 54) = 5.783, p < .05$. Performance ratings were higher when both the supervisor and the subordinate group provided information encouraging high performance ratings ($M = 4.20, SD = .67$) than low performance ratings ($M = 3.77, SD = 0.68$). Tukey's test also suggests that the difference between the two means (0.43) was significantly different. Therefore, the results provide support for Hypothesis 6. A bar chart displaying this effect is illustrated in Figure 5 on the following page.

Figure 5

Effects of Supervisor and Subordinate Group Deflation Versus Inflation Information on Performance Ratings (Hypothesis 6)

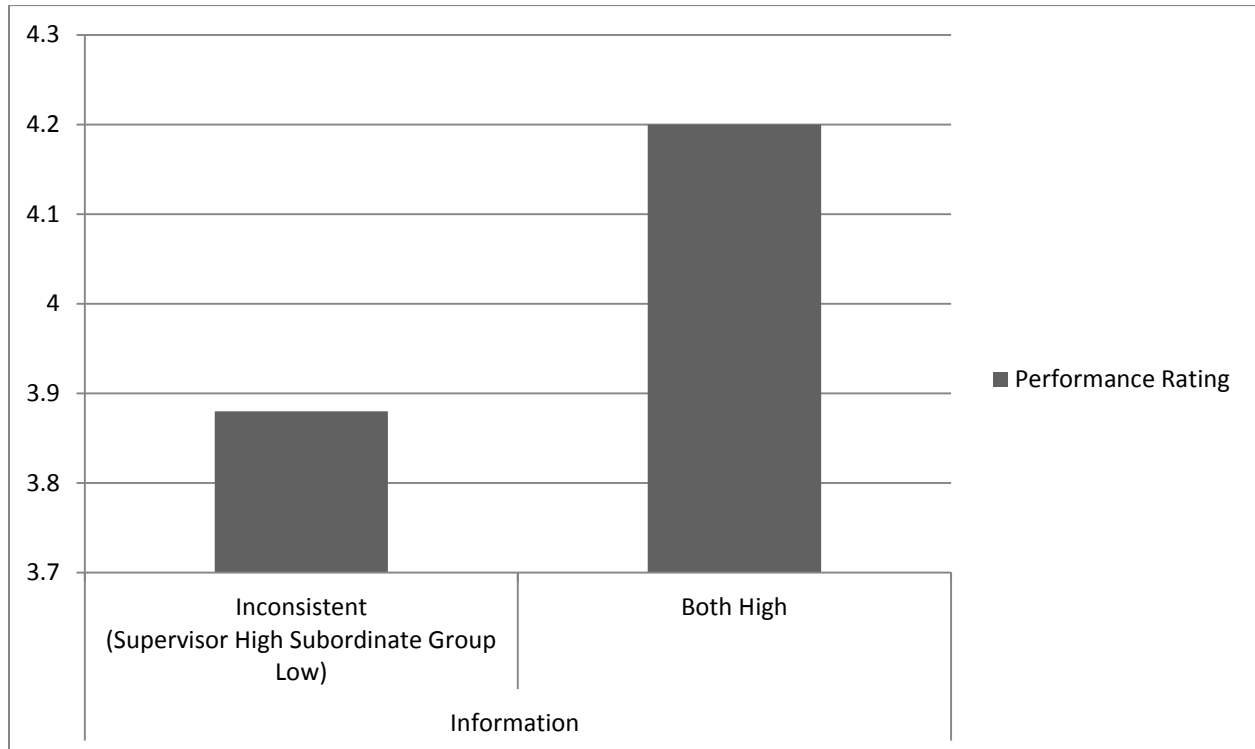


Hypothesis 7 predicted that supervisor and subordinate group inflation information will have a stronger effect on performance ratings than supervisor and subordinate group deflation information. The χ^2 test indicates that the effect magnitude of supervisor and subordinate group inflation information on performance ratings was not significantly different from the effect magnitude of supervisor and subordinate group deflation information on performance ratings, $\chi^2(1) = 2.31, ns$. Therefore, Hypothesis 7 was not supported.

Hypothesis 8a posits that performance ratings will be higher when both the subordinate group and supervisor provide information encouraging high performance ratings than when the supervisor provides inflation information and the subordinate group provides deflation information. To test Hypothesis 8a, a One-way ANOVA was conducted with direction of inconsistent information (supervisor and subordinate group information encouraging high performance ratings, supervisor encouraging high performance ratings and subordinate group encouraging low performance ratings) as the between-subjects factor. Results indicate the direction of inconsistent information had a significant effect on performance ratings, $F(1, 55) = 3.005, p < .05$ (one-tailed test), as performance ratings were higher when both the supervisor and subordinate group provided information encouraging high performance ratings ($M = 4.20, SD = .67$) than when the supervisor provided information encouraging high performance ratings and the subordinate group provided information encouraging low performance ratings ($M = 3.88, SD = .75$). Tukey's test suggests that the difference between the two means (0.32) was not significantly different, however. Therefore, Hypothesis 8a was supported by the results of the One-way ANOVA, but not Tukey's test. A bar chart displaying this effect is illustrated in Figure 6 on the following page.

Figure 6

Effects of Inconsistent Information Versus Supervisor and Subordinate Group Inflation
Information on Performance Ratings (Hypothesis 8a)

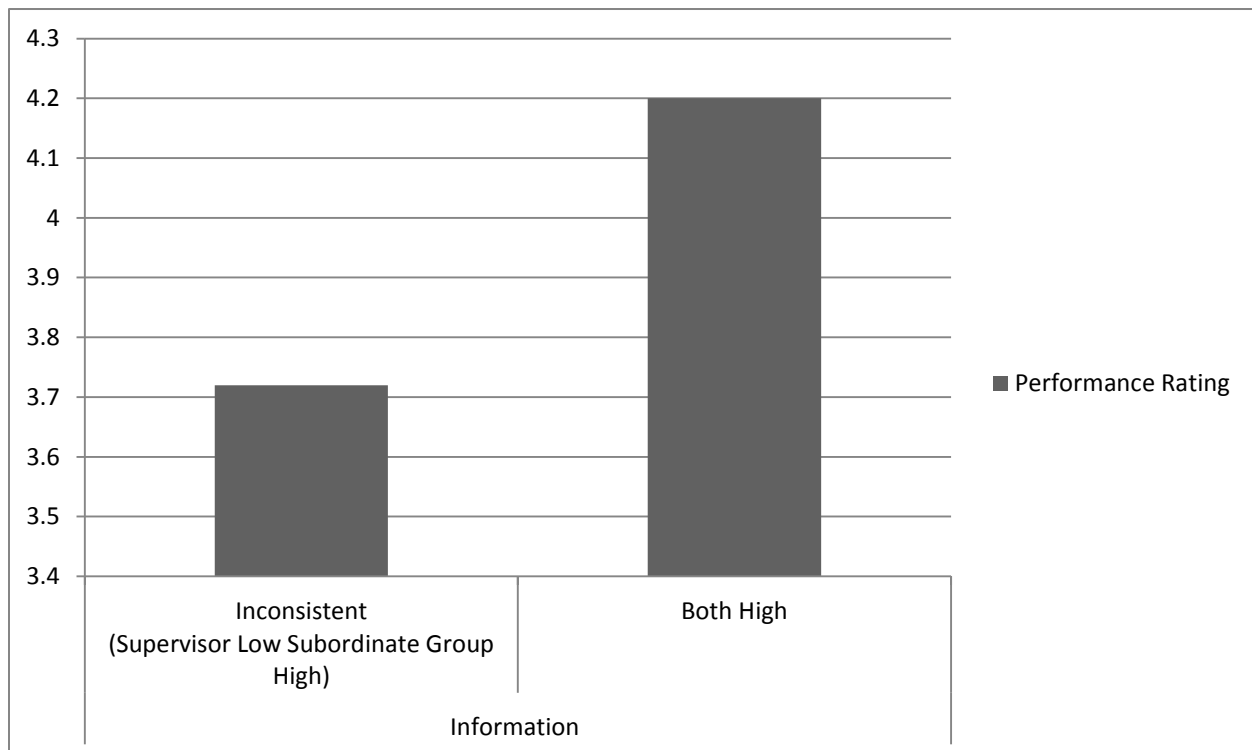


In Hypothesis 8b, it was predicted that performance ratings will be higher when the subordinate group and supervisor both provide information encouraging high performance ratings than when the supervisor provides deflation information and the subordinate group provides inflation information. To test Hypothesis 8b, a One-way ANOVA with direction of inconsistent information (supervisor encourage low performance ratings and subordinate group encourages high performance ratings, both provide information encouraging high performance ratings) as the between-subjects factor. Results indicate that the direction of inconsistent

information had an effect on performance ratings, $F(1, 54) = 7.667$, $p < .01$. Performance ratings were significantly higher in the consistent information condition ($M = 4.20$, $SD = .67$) than in the inconsistent performance ratings condition ($M = 3.72$, $SD = .65$). Tukey's test also suggests that the difference between the two means (0.48) was significantly different. As such, Hypothesis 8b was also supported. A bar chart displaying this effect is illustrated in Figure 7 below.

Figure 7

Effects of Inconsistent Information Versus Supervisor and Group Inflation Information on Performance Ratings (Hypothesis 8b)



Hypothesis 9a stated that performance ratings will be lower when both the supervisor and subordinate group provide information encouraging low performance ratings than when the supervisor provides information encouraging high performance ratings and the subordinate group provides information encouraging low performance ratings. Hypothesis 9a was tested using a One-way ANOVA, with consistency of information (supervisor encourages high ratings and subordinate group encourage low ratings, both encouraging low performance ratings) as the between-subjects factor. Results indicate there was no significant effect of consistency of information on performance ratings, $F(1, 57) = .331, ns$, as performance ratings were not significantly different when both the supervisor and subordinate group encouraged low performance ratings ($M = 3.77, SD = .68$) than when the supervisor provided information encouraging high ratings and the teller was perceived to be a bad employee by the subordinate group ($M = 3.88, SD = .75$). Tukey's test also suggests that the difference between the two means (0.11) was not significantly different. Therefore, Hypothesis 9a failed to receive support.

Hypothesis 9b stated that performance ratings will be lower when both the supervisor and subordinate group provide information encouraging low performance ratings than when the supervisor provides information encouraging low performance ratings and the subordinate group provides information encouraging high performance ratings. Hypothesis 9b was also tested using a One-way ANOVA with consistency of information (supervisor encourages low performance ratings and subordinate group encourage high performance ratings, both encourage low performance ratings) as the between-subjects factor. Results indicate there was no significant effect on performance ratings, $F(1, 56) = .087, ns$. Performance ratings were not significantly different in the inconsistent information condition ($M = 3.72, SD = .65$) than in the consistent information condition ($M = 3.77, SD = .68$). Tukey's test also suggests that the difference

between the two means (0.05) was not significantly different. Hypothesis 9b was therefore not supported.

Hypothesis 10a predicted that performance ratings will be higher when the supervisor provides information encouraging high performance ratings and the subordinate group provides information encouraging low performance ratings than when the subordinate group provides information encouraging high performance ratings and the supervisor provides information encouraging low performance ratings. Hypothesis 10b posited that performance ratings will be lower when the supervisor provides information encouraging low performance ratings and the subordinate group provides information encouraging high performance ratings than when the supervisor provides information encouraging high performance ratings and the subordinate group provides information encouraging low performance ratings. To test these two hypotheses, the same group means are compared.

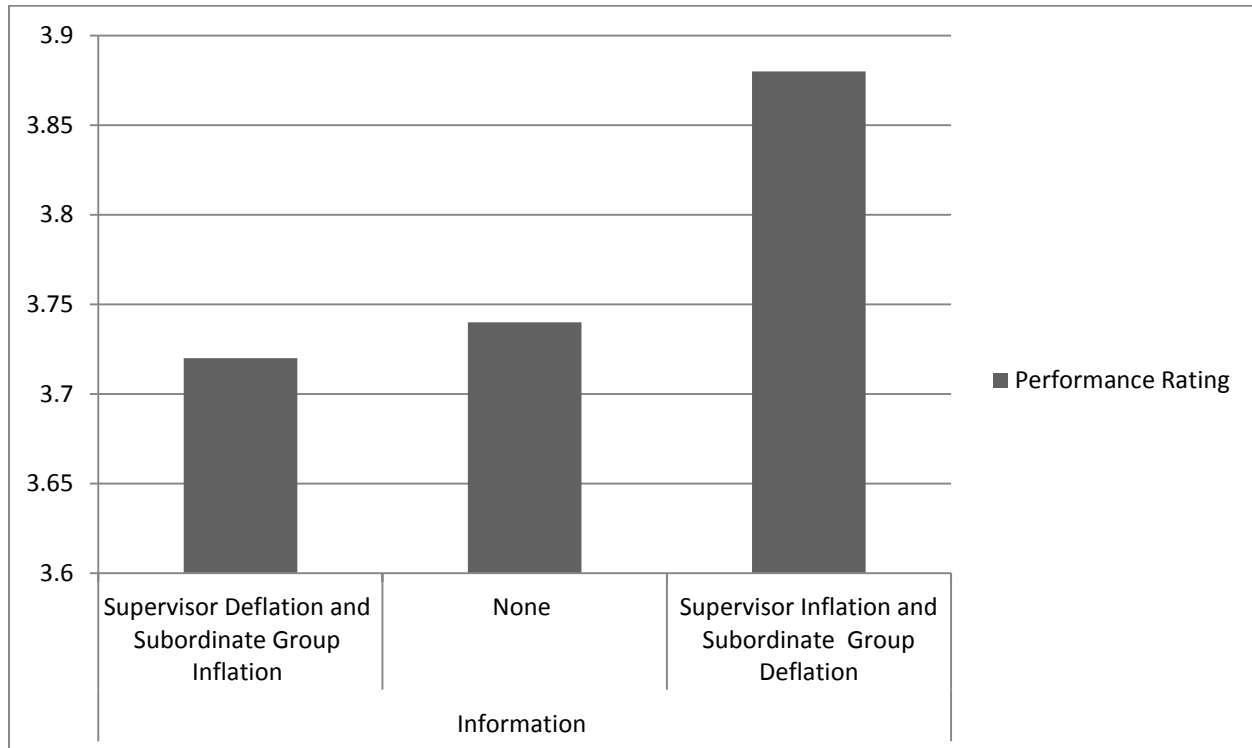
Therefore, a One-way ANOVA was conducted with direction of inconsistent information (supervisor encourages high ratings and subordinate group encourages low ratings, supervisor encourages low ratings and subordinate group encourages high ratings) as the between-subjects factor, to test Hypotheses 10a and 10b. The results indicate that the direction of inconsistent information did not have a significant effect on performance ratings, $F(1, 57) = .762, ns$. While not significantly different, ratings were higher when the supervisor encouraged high performance ratings and subordinate group encouraged low performance ratings ($M = 3.88, SD = 0.75$) than when the supervisor encouraged low performance ratings and subordinate group encouraged high performance ratings ($M = 3.72, SD = 0.65$). Tukey's test also suggests that the difference between the two means (0.16) was not significantly different. Hypothesis 10a and 10b were not supported.

Hypothesis 10c predicted that supervisor inflation information and subordinate group deflation information will have a stronger effect on performance ratings than supervisor deflation information and subordinate group inflation information. To obtain effect sizes, two One-way ANOVAs were conducted, one with supervisor inflation and group deflation information (supervisor inflation information and subordinate group deflation information, no information) and one with supervisor deflation information and subordinate group inflation information (supervisor deflation information and group inflation information, no information).⁸ The χ^2 test indicates that the effect magnitude of inflation information on performance ratings was significantly different from the effect magnitude of deflation information on performance ratings, $\chi^2(1) = 3.33, p < .05$ (one-tailed test). Therefore, Hypothesis 10c received support. A bar chart of these effects is depicted in Figure 8 on the following page.

⁸ Performance ratings in the supervisor inflation and subordinate group deflation information condition ($M = 3.87, SD = 0.75$) were not significantly different, $F(1, 58) = .428, ns, \eta^2 = .007$, from ratings in the no information condition ($M = 3.74, SD = .83$). The results indicate that performance ratings in the supervisor deflation information and subordinate group inflation information condition ($M = 3.72, SD = 0.65$) were not significantly different, $F(1, 58) = .018, ns, \eta^2 = .000$, from the ratings in the no information condition ($M = 3.74, SD = .83$).

Figure 8

Effect of Supervisor Inflation and Subordinate Group Deflation Information Versus Supervisor Deflation and Subordinate Group Inflation Information on Performance Ratings (Hypothesis 10c)



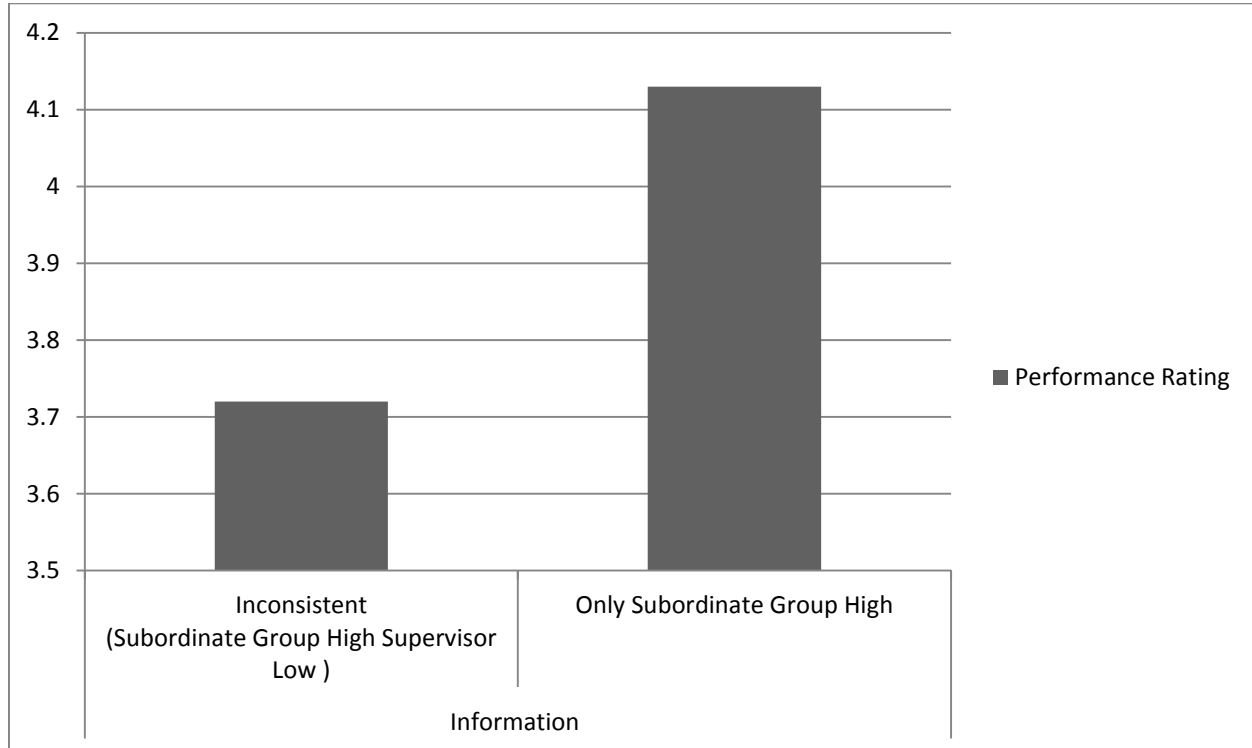
In Hypothesis 11a, I predicted that performance ratings will be higher when only the supervisor provides information encouraging high performance ratings than when the supervisor provides information encouraging high performance ratings and the subordinate group provides information encouraging low performance ratings. A One-way ANOVA was performed to test Hypothesis 11a with information (only supervisor encourages high performance ratings, supervisor encourages high performance ratings and subordinate group encourages low

performance ratings) as the between-subjects factor. Results indicate there was no significant effect on performance ratings, $F(1, 57) = .012, ns$. Performance ratings were not significantly different in the supervisor encourages high performance ratings and subordinate group encourages low performance ratings condition ($M = 3.88, SD = .75$) than in the only the supervisor encourages high performance ratings condition ($M = 3.90, SD = .74$). Tukey's test also suggests that the difference between the two means (0.02) was not significantly different. Hypothesis 11a did not receive support.

Hypothesis 11b predicted that performance ratings will be higher when only the subordinate group conveys information encouraging high performance ratings than when the subordinate group provides information encouraging high performance ratings and the supervisor provides information encouraging low performance ratings. A One-way ANOVA was performed to test Hypothesis 11b with information (supervisor encourages low performance ratings and subordinate group encourages high performance ratings, only subordinate group encourages high performance ratings) as the between-subjects factor. Results indicate there was a significant effect on performance ratings, $F(1, 56) = 4.667, p < .05$. Performance ratings were lower when the supervisor encourages low performance ratings and subordinate group encourages high performance ratings ($M = 3.72, SD = .65$) than when only the subordinate group's information encourages high performance ratings ($M = 4.13, SD = .81$). Tukey's test also suggests that the difference between the two means (0.41) was significantly different. Hypothesis 11b received support. A bar chart displaying this effect is illustrated in Figure 9 on the following page.

Figure 9

Effects of Inconsistent Information Versus Supervisor and Subordinate Group Inflation Information on Performance Ratings (Hypothesis 11b)



In Hypothesis 11c, it was predicted that performance ratings will be lower when only the supervisor provides information encouraging low performance ratings than when the supervisor provides information encouraging low performance ratings and the subordinate group conveys information encouraging high performance ratings. A One-way ANOVA, with source of information (only supervisor provides deflation information, supervisor provides deflation information and subordinate group provides inflation information) as the between-subjects factor, was conducted to test Hypothesis 11c. Results indicate there was no significant effect on performance ratings, $F(1, 57) = 1.815, ns$. Performance ratings were not significantly different when inconsistent information was presented ($M = 3.72, SD = .65$) than when only supervisor

deflation information was presented ($M = 3.98, SD = .86$). Tukey's test also suggests that the difference between the two means (0.26) was not significantly different. Hypothesis 11c was therefore not supported.

In Hypothesis 11d, I predicted that ratings will be lower when only the subordinate group provides information encouraging low performance ratings than when the subordinate group provides information encouraging low performance ratings and the supervisor provides information encouraging high performance ratings. To test Hypothesis 11d, a One-way ANOVA was conducted with source of information (only subordinate group provides deflation information, subordinate group provides deflation information and supervisor provides inflation information) as the between-subjects factor. Results indicate there was no significant effect on performance ratings, $F(1, 57) = .867, ns$. Performance ratings were not significantly different when inconsistent information was presented ($M = 3.89, SD = .75$) than when only subordinate group deflation information was presented ($M = 3.98, SD = .86$). Tukey's test also suggests that the difference between the two means (0.09) was not significantly different. Hypothesis 11d was therefore not supported.

Hypothesis 12a predicted that performance ratings will be higher when the supervisor provides information encouraging high performance ratings and the subordinate group provides information encouraging low performance ratings than when no information is provided. Hypothesis 12a was tested using a One-way ANOVA with information (supervisor encourages high performance ratings and subordinate group encourages low performance ratings, no information) as the between-subjects factor. Results indicate there was no significant effect on performance ratings, $F(1, 58) = .428, ns$. Performance ratings were not significantly different in the inconsistent information condition ($M = 3.88, SD = .75$) than in the no information condition

($M = 3.74$, $SD = .83$). Tukey's test also suggests that the difference between the two means (0.14) was not significantly different. Hypothesis 12a was therefore not supported.

Hypothesis 12b posited that performance ratings will be lower when the subordinate group provides information encouraging high performance ratings and the supervisor provides information encouraging low performance ratings than when no information is provided. A One-way ANOVA was performed to test Hypothesis 12b with information (supervisor encourages low performance ratings and subordinate group encourages high performance ratings, no rating information) as the between-subjects factor. Results indicate there was no significant effect on performance ratings, $F(1, 57) = .018$, *ns*. Performance ratings were not significantly different in the supervisor encourages low performance ratings and subordinate group encourages high performance ratings condition ($M = 3.72$, $SD = .65$) than in the no rating information condition ($M = 3.74$, $SD = .83$). Tukey's test also suggests that the difference between the two means (0.026) was not significantly different. Hypothesis 12b did not receive support.

Follow-Up Analyses

Effects of Manipulations on Desire to Inflate or Deflate Ratings

Given the focus on rater motivation and the number of unsupported hypotheses, follow-up analyses were conducted. Specifically, the extent to which participants wanted to provide distorted performance ratings was assessed. Three items were used to measure the extent to which participants wanted to provide high performance ratings for the bank teller (i.e., Julie). These include: "My life will be easier if I give Julie a high performance rating," "I want to give Julie a good evaluation," "I think Julie should get a high performance rating." Three items (i.e., "I am motivated to give Julie very low ratings," "I should give Julie low ratings," "All in all, I

think Julie should get a low evaluation”) were used to assess the extent to which participants wanted to provide Julie with low performance ratings. Participants responded to all items using a 1 (strongly disagree) to 5 (strongly agree) scale.

Results of a univariate ANOVA suggest that the participants’ desire to provide *high* performance ratings was affected by the subordinate group manipulation [$F(2, 261) = 53.74, p < .001$] and the supervisor manipulation [$F(2, 261) = 9.51, p < .001$]. The interaction effect of the subordinate group and supervisor on the participants’ want to provide inflated performance ratings, however, was not significant, $F(4, 261) = .96, ns$. Tukey’s follow-up test indicates that the means were significantly different across all subordinate group manipulations. Means were not significantly different between the supervisor no information and supervisor low ratings information conditions. Means for the extent to which participants wanted to provide high performance ratings across the different manipulations are shown in Table 11 below.

Table 11

Means for Participants’ Desire to Provide High Performance Ratings Across Experimental Conditions

Group Manipulation	Supervisor Manipulation		
	None	Low	High
None	3.08	3.02	3.51
Low	2.33	2.58	2.64
High	3.26	3.36	3.66

A second univariate ANOVA was conducted to determine if the extent to which participants wanted to provide *low* performance ratings was affected by the subordinate group and supervisor manipulations. Results indicate that the subordinate group manipulation had a significant main effect on participants' desire to provide low performance ratings, $F(2, 261) = 74.48, p < .001$. The supervisor manipulation did not have a significant main effect on participants' want to provide low performance ratings, $F(2, 261) = 1.21, ns$. There was a significant interaction effect of the subordinate group and supervisor manipulations on the participants' want to provide low performance ratings, $F(4, 261) = 3.43, p < .01$. Means for the extent to which participants wanted to provide low performance ratings across the different manipulations are shown in Table 12 below.

Table 12

Means for Participants' Desire to Provide Low Performance Ratings Across Experimental Conditions

Subordinate Group Manipulation	Supervisor Manipulation		
	None	Low	High
None	2.20	2.41	2.08
Low	3.30	2.68	3.29
High	1.73	1.63	1.74

These follow-up tests indicate that the supervisor and subordinate group manipulations had some effect on participants' desire to provide high and low performance ratings for the bank

teller. Specifically, participants wanted to provide high ratings to a greater extent when the subordinate group or supervisor provided information encouraging high performance ratings. There was no interaction effect of supervisor and subordinate group information on the participants' want to provide high performance ratings, however. Additionally, while the subordinate group had an effect on participants want to provide low performance ratings, the supervisor did not have an effect. The interaction effect of the supervisor and subordinate group was significant. Interestingly, raters wanted to provide low performance ratings to a lesser extent when both the subordinate group and supervisor provided information encouraging low performance ratings than when the subordinate group provided information encouraging low performance ratings and the supervisor either provided (a) no information or (b) information encouraging high performance ratings.

To examine the effects of these manipulations in greater detail on rater motivation, additional follow up tests were conducted. The first set of tests focused on $P \rightarrow O$ expectancies, which were predicted to be affected by the context. The context may also have effects on state valences, such that supervisor approval of performance ratings is valued to a greater extent when it is more salient (i.e., when the supervisor expresses a preference). Examining the effects of the social contextual factors on $P \rightarrow O$ expectancies and state valences allows us to determine if and how the type of information (i.e., inflation versus deflation information) influences state valences of supervisor and subordinate group approval of performance ratings and the extent to which accurate ratings are perceived to result in approval.

Effects of Manipulations on $P \rightarrow O$ Expectancies

$P \rightarrow O$ expectancies that accurate performance ratings will result in approval from the supervisor were measured using five items. These include: “The accuracy of my ratings will not affect how my boss feels about me (reverse-scored),” “How happy my boss is depends on how accurately I rate the teller’s performance,” “If I rate accurately, my boss will be happier with me,” “The more accurately I rate, the happier my boss will be,” “My boss will be happy [if I rate performance accurately].” Five items were also used to measure $P \rightarrow O$ expectancies that accurate ratings will result in subordinate group approval. These include: “The accuracy of my ratings will not affect how my coworkers feel about me (reverse-scored),” “The more accurately I rate, the happier my coworkers will be,” “How happy my coworkers are depends on how I accurately I rate the teller’s performance,” “If I rate accurately, my coworkers will be happier with me,” and “My group will be happy [if I rate performance accurately].” Participants responded to all items using a 1 (strongly disagree) to 5 (strongly agree) scale.

The type of information may have an effect on raters’ $P \rightarrow O$ expectancies. Therefore, two One-way ANOVAs were performed to assess participants’ $P \rightarrow O$ expectancies. The first One-way ANOVA was performed with supervisor information (no rating information, high ratings, low ratings) as the between-subjects factor. Results show a significant effect of the supervisor information manipulation on the extent to which participants believed that accurate ratings would result in supervisor approval, $F(2, 267) = 17.80, p < .001$. Tukey’s test indicates that participants were less likely to perceive a link between providing accurate performance ratings and supervisor approval when the supervisor provided information encouraging low performance ratings ($M = 3.31, SD = .68$) than when the supervisor provided no rating information ($M = 3.78, SD = .57$) or provided information encouraging high performance ratings

($M = 3.77$, $SD = .56$). This indicates that participants thought that providing accurate ratings would be less likely to result in approval when the supervisor preferred low performance ratings for the teller than when the supervisor provided no information. Providing accurate ratings was not perceived to affect supervisor approval when the supervisor provided information encouraging high performance ratings (as there was no difference between the no rating information and high rating conditions).

Another One-way ANOVA was performed with subordinate group rating preference (no rating preference, good employee, bad employee) as the between-subjects factor. Results indicate that subordinate group information did not have an effect on the extent to which participants believed that accurate ratings would result in subordinate group approval, $F(2, 267) = .36$, *ns*. The $P \rightarrow O$ expectancy for accurate ratings resulting in subordinate group approval was not different when the subordinate group provided no information ($M = 2.57$, $SD = .58$) versus when the subordinate group thought the ratee was a good employee ($M = 2.50$, $SD = .74$) or a bad employee ($M = 2.55$, $SD = .57$). This suggests that participants did not think that the rating they provided for the bank teller would affect subordinate group approval of ratings.

Effects of Manipulations on Valences

The type of information raters encounter may also affect the extent to which they value the supervisor's and subordinate group's approval of the performance ratings they render. To investigate this, the effects of the manipulations on valences were examined. Valences for supervisor approval were measured using two items (i.e., "I want to make my boss happy with my ratings," "I want my boss to like my ratings"). Valences for subordinate group approval were also measured using two items (i.e., "I want my coworkers to like my ratings," "I want to make

my coworkers happy with my ratings”). Participants responded to all items using a 1 (strongly disagree) to 5 (strongly agree) scale.

A One-way ANOVA was performed with supervisor information (no rating information, high ratings, low ratings) as the between-subjects factor. Results show a significant effect of the supervisor information manipulation on the extent to which supervisor approval was valued, $F(2, 267) = 3.00, p \leq .05$. Tukey’s test indicates that supervisor approval of performance ratings was valued more when the supervisor provided information encouraging high performance ratings ($M = 3.81, SD = .73$) than low performance ratings ($M = 3.54, SD = .80$). There were no differences between these two conditions and the no supervisor information condition ($M = 3.71, SD = .68$). These results suggest that supervisor approval of ratings is valued more when the supervisor expresses a preference for high performance ratings than when the supervisor expresses a preference for low performance ratings.

Another One-way ANOVA was performed with subordinate group information (no ratee information, bad ratee, good ratee) as the between subjects factor. Results indicate that subordinate group information did not have an effect on the extent to which subordinate group approval of performance ratings was valued, $F(2, 267) = .62, ns$. This suggests that the extent to which subordinate group approval of ratings is valued is not affected by the type of information the subordinate group provides about a ratee or even whether the group expresses an opinion at all.

Taken together, the follow-up tests indicate that participants valued supervisor approval of performance ratings more when the supervisor provided information encouraging high performance ratings than when the supervisor provided information encouraging low

performance ratings. The extent to which subordinate group approval of ratings was valued did not change based on whether, or the type of information the group provided. Finally, participants did not perceive that the extent to which they provide accurate performance ratings affects subordinate group approval of ratings. They did, however, perceive that they were less likely to attain supervisor approval when the supervisor provided information encouraging low performance ratings. As with the subordinate group, when the supervisor provided information supporting high performance ratings, raters did not perceive that accurate ratings were less likely to result in supervisor approval.

Interaction Effects of Manipulations

The hypotheses that were supported in the presented study largely dealt with interaction effects between supervisor and subordinate group information. Therefore, a number of univariate ANOVAs were conducted to examine the extent to which information from the supervisor and subordinate group interacts to influence raters' valences of supervisor and subordinate group approval of performance ratings and $P \rightarrow O$ expectancies that accurate performance ratings will result in supervisor and subordinate group approval.

Interaction Effects of Manipulations on Valences

Results indicate that information from the supervisor and subordinate group⁹ interacted to affect valence of supervisor approval of performance ratings, $F(4,261) = 3.291, p < .05$. For example, supervisor approval of performance ratings had a higher valence when both the subordinate group and supervisor provided information encouraging high performance ratings ($M = 4.00, SD = .79$) than when the supervisor encouraged low performance ratings and the

⁹ There was no main effect of the subordinate group on the valence of supervisor approval of performance ratings, $F(2,261) = .070, ns$.

group encouraged high performance ratings ($M = 3.28$, $SD = .96$). Table 13, below, contains means for the valence of supervisor approval of performance ratings across all experimental conditions. The interaction of supervisor and subordinate group information did not have a significant effect on the valence of group approval of performance ratings, $F(4,261) = .174$, *ns*. Table 14, on the next page, contains means for the valence of subordinate group approval of performance ratings across experimental conditions.

Table 13

Means for Participants' Valence of Supervisor Approval of Performance Ratings Across Experimental Conditions

Subordinate Group Manipulation	Supervisor Manipulation		
	None	Low	High
None	3.55	3.58	3.83
Low	3.80	3.75	3.58
High	3.78	3.28	4.00

Table 14

Means for Participants' Valence of Subordinate Group Approval of Performance Ratings Across Experimental Conditions

Subordinate Group Manipulation	Supervisor Manipulation		
	None	Low	High
None	3.22	3.08	3.30
Low	3.13	3.20	3.38
High	3.12	3.03	3.17

Interaction Effects of Manipulations on $P \rightarrow O$ Expectancies

The interaction of supervisor and subordinate group information also had an effect on participants' $P \rightarrow O$ expectancies that accurate performance ratings result in supervisor approval of performance ratings, $F(4,261) = 6.346, p < .01$ ¹⁰. For example, raters' $P \rightarrow O$ expectancy that accurate performance ratings result in supervisor approval of the ratings was lower when the supervisor provided information encouraging low performance ratings and the subordinate group provided information encouraging high performance ratings ($M = 2.99, SD = .68$) than when both provided information encouraging high performance ratings ($M = 3.99, SD = .62$) or when no information was provided ($M = 3.89, SD = .50$). Table 15, on the following page, contains means, across experimental conditions, for the $P \rightarrow O$ expectancies that accurate performance

¹⁰ There was no main effect of the subordinate group on the $P \rightarrow O$ expectancy regarding the extent to which accurate performance ratings result in supervisor approval, $F(2,261) = .622, ns$.

ratings will result in supervisor approval of performance ratings. The interaction of supervisor and subordinate group information did not have a significant effect on participants' $P \rightarrow O$ expectancies that accurate performance ratings result in group approval of performance ratings, $F(4,261) = 1.202, ns$. Table 16, on the following page, contains means, across experimental conditions, for the $P \rightarrow O$ expectancies that accurate performance ratings will result in subordinate group approval of performance ratings.

Table 15

Means for Participants' $P \rightarrow O$ Expectancies of Accurate Performance Ratings Resulting in Supervisor Approval of Performance Ratings

Across Experimental Conditions

Subordinate Group Manipulation	Supervisor Manipulation		
	None	Low	High
None	3.89	3.35	3.79
Low	3.69	3.59	3.54
High	3.76	2.99	3.99

Table 16

Means for Participants' $P \rightarrow O$ Expectancies of Accurate Performance Ratings Resulting in Subordinate Group Approval of Performance Ratings Across Experimental Conditions

Subordinate Group Manipulation	Supervisor Manipulation		
	None	Low	High
None	3.44	3.27	3.57
Low	3.46	3.39	3.51
High	3.29	3.57	3.65

The last set of follow-up analyses indicates that supervisor and subordinate group information interacted to affect both valences and $P \rightarrow O$ expectancies associated with supervisor approval but not those associated with subordinate group approval. This suggests that information from the subordinate group had an effect on the extent to which supervisor approval was valued and the extent to which participants perceived that accurate ratings will result in supervisor approval.

DISCUSSION

The performance appraisal context is increasingly being taken into consideration by researchers (e.g., Murphy & Cleveland, 1995). The social context of performance appraisal, however, while important, has received inadequate research attention (Levy & Williams, 2004). The limited studies that have been conducted (e.g., Judge & Ferris, 1993; Mero et al., 2007; Spence & Keeping, 2010) provide evidence that social contextual factors affect performance ratings. Spence and Keeping (2010) have suggested that in order to offer practical recommendations for improving performance appraisal accuracy, we first need to understand the effects of the social context on performance ratings. This study examined the effects of two key, yet largely unexamined, social contextual factors (i.e., the rater's supervisor and subordinate group) on performance ratings.

Rater goal theory and findings in the performance appraisal literature (e.g., Longenecker et al., 1987; Wang et al., 2010) suggest that rating is a motivated behavior. Therefore, a motivational approach to understanding the effects of these key social contextual factors was taken. Specifically, expectancy theory was used as it allows fine-grained examination of how various motivational components are affected by contextual factors. This diagnostic approach can be used to provide practical recommendations (Sanchez et al., 2000) as we gain an understanding of what motivational components are affected by the social influences and how. To supplement expectancy theory, and inform our understanding of the valences associated with the supervisor and the group when both are salient sources of information in performance appraisal, social impact theory and the social influence literature were integrated with the expectancy framework.

Overall, the results provide some evidence that key factors in the social context of performance appraisal (i.e., the rater's supervisor and subordinate group) influence performance ratings. Specifically, when both the supervisor and subordinate group expressed a preference for high performance ratings, ratings were higher than when (a) no information, (b) inconsistent information, or (c) deflation information were encountered. Additionally, performance ratings were higher when the subordinate group provided inflation information and the supervisor provided no information than when (a) the subordinate group provided inflation and the supervisor provided deflation information or (b) no information was provided. A summary of the hypotheses and results is displayed in Table 17 on the following two pages.

These findings illustrate that the social context of performance appraisal influences performance ratings and provide empirical support for rater goal theory (e.g., Cleveland & Murphy, 1992). Furthermore, the results suggest that the goals raters have, which have been shown to influence performance ratings (e.g., Wang et al., 2010; Wong & Kwong, 2007), are affected by the social context in which the performance appraisal process transpires. The findings also indicate that inflation information has a stronger effect on performance ratings than deflation information. In other words, raters are more likely to provide ratings consistent with information provided by the supervisor and subordinate group when the information encourages high, compared to low, performance ratings. Finally, the pattern of results indicates that it is important to consider the complexities inherent in performance appraisal and avoid simplistic conceptualizations of the performance appraisal context and its effects.

Table 17

Summary of Hypotheses and Results

Hypothesis 1: Inflation information will have a stronger effect on performance ratings than deflation information.	Supported
Hypothesis 2a: Performance ratings will be lower when the supervisor provides deflation information than when no information is provided.	Not Supported
Hypothesis 2b: Performance ratings will be higher when the supervisor provides inflation information than when no information is provided.	Not Supported
Hypothesis 2c: Supervisor inflation information will have a stronger effect on performance ratings than supervisor deflation information.	Not Supported
Hypothesis 3a: Performance ratings will be lower when the subordinate group provides deflation information than when no information is provided.	Not Supported
Hypothesis 3b: Performance ratings will be higher when the subordinate group provides inflation information than when no information is provided.	Supported
Hypothesis 3c: Subordinate group inflation information will have a stronger effect on performance ratings than subordinate group deflation information.	Supported
Hypothesis 4: Performance ratings will be higher when both the supervisor and subordinate group provide consistent inflation information than when (a) only the supervisor provides inflation information, (b) only the subordinate group provides inflation information, or (c) no information is provided.	Only H4c was Supported
Hypothesis 5: Performance ratings will be lower when both the supervisor and subordinate group provide consistent deflation information than when (a) only the supervisor provides deflation information, (b) only the subordinate group provides deflation information, or (c) no information is provided.	Not Supported
Hypothesis 6: Performance ratings will be higher when both the supervisor and subordinate group provide consistent inflation information than deflation information.	Supported
Hypothesis 7: Subordinate group and supervisor inflation information will have a stronger effect on performance ratings than subordinate group and supervisor deflation information.	Not Supported
Hypothesis 8: Performance ratings will be higher when both the subordinate group and supervisor provide inflation information than when (a) the supervisor provides inflation information and the subordinate group provides deflation information, or (b) when the supervisor provides deflation information and the subordinate group provides inflation information.	Supported
Hypothesis 9: Performance ratings will be lower when both the subordinate group and supervisor provide deflation information than when (a) the supervisor provides inflation information and the subordinate group provides deflation information, or (b) when the supervisor provides deflation information and the subordinate group provides inflation information.	Not Supported
Hypothesis 10a: Performance ratings will be higher when the rater's supervisor provides inflation information and the subordinate group provides deflation information than when the subordinate group provides inflation information and the supervisor provides deflation information.	Not Supported

Hypothesis 10b: Performance ratings will be lower when the rater's supervisor provides deflation information and the subordinate group provides inflation information than when the subordinate group provides deflation information and the supervisor provides inflation information.	Not Supported
Hypothesis 10c: Supervisor inflation information and subordinate group deflation information will have a stronger effect on performance ratings than supervisor deflation information and subordinate group inflation information.	Supported
Hypothesis 11a: Performance ratings will be higher when only the supervisor provides inflation information than when the supervisor provides inflation information and the subordinate group provides deflation information.	Not Supported
Hypothesis 11b: Performance ratings will be higher when only the subordinate group provides inflation information than when the subordinate group provides inflation information and the supervisor provides deflation information.	Supported
Hypothesis 11c: Performance ratings will be lower when only the supervisor provides deflation information than when the supervisor provides deflation information and the subordinate group provides inflation information.	Not Supported
Hypothesis 11d: Performance ratings will be lower when only the subordinate group provides deflation information than when the subordinate group provides deflation information and the supervisor provides inflation information.	Not Supported
Hypothesis 12a: Performance ratings will be higher when the supervisor provides inflation information and the subordinate group provides deflation information than when no information is provided.	Not Supported
Hypothesis 12b: Performance ratings will be lower when the supervisor provides deflation information and the subordinate group provides inflation information than when no information is provided.	Not Supported

The limited number of studies (e.g., Mero et al., 2007; Spence & Keeping, 2010) that have focused on the effects of the social context on performance ratings had limitations that were addressed in the present study. Spence and Keeping (2010) for example, presented objective performance information, which, as they indicate, is not likely to be the case in organizations. Given the objectively presented performance information, social contextual factors may not affect performance ratings in the same way as they may in organizational settings (Spence & Keeping, 2010), where performance information has to be actively observed and judged. Mero and colleagues (2007) focused on accountability to different audiences, but the rating preferences

of their audiences were unknown to raters. To address the limitations of previous research, and simulate a more realistic environment, two steps were taken in the present study. One, raters had to observe, evaluate, and report the performance of the ratee. Two, the preferences of key social contextual factors, the rater's supervisor and subordinate group, were made more explicit.

Summary of Supervisor Main Effects

Contrary to expectations, inflation and deflation information from the supervisor did not have a main effect on performance ratings. While not significant, ratings were higher when the supervisor provided information encouraging high performance ratings as well as low performance ratings than when no information was provided from either source (i.e., the control condition). The follow-up analyses that were conducted suggest that this may have occurred because the motivational force to provide ratings consistent with the supervisor's information was not sufficiently high. The valence of supervisor approval was not significantly different when the supervisor encouraged high performance ratings compared to when the supervisor provided no rating information. Additionally, the $P \rightarrow O$ expectancy that accurate performance ratings will result in supervisor approval of performance ratings was not lower when the supervisor expressed a preference for high performance ratings than when no preference was conveyed. Therefore, raters' subjective probability estimates that providing inflated performance ratings will result in supervisor approval and the extent to which supervisor approval of the ratings is valued was not different when the supervisor provided no information than when the supervisor provided information encouraging high performance ratings. Taken together, supervisor inflation information did not have a significant main effect on performance ratings because the supervisor information did not affect the valence of supervisor approval of ratings or

the raters' $P \rightarrow O$ expectancies regarding the extent to which accurate performance ratings result in supervisor approval of ratings.

When the supervisor provided deflation information, results indicate that the $P \rightarrow O$ expectancy that accurate ratings will result in supervisor approval of ratings was lower than when no supervisor information was provided. This indicates that participants perceived that accurate ratings were less likely to result in supervisor approval when the supervisor expressed a preference for low performance ratings than when no supervisor information was communicated. The valence of supervisor approval was not different when the supervisor expressed a preference for low performance ratings compared to when no supervisor information was provided. The follow-up analyses indicate that the supervisor manipulation did not have an effect on raters' desire to provide low performance ratings, such that the extent to which participants wanted to provide low performance ratings for the teller did not differ across these two conditions. This suggests that the motivational force to provide deflated performance ratings for the teller when the supervisor encouraged low performance ratings may not have been sufficiently strong to outweigh the motivational force to provide accurate performance ratings.

Summary of Subordinate Group Main Effects

Consistent with predictions, performance ratings were higher when the subordinate group provided information encouraging high performance ratings than when no information was provided. The follow-up analyses focusing on rater motivation demonstrate that $P \rightarrow O$ expectancies that accurate performance ratings will result in subordinate group approval of ratings were not significantly different when the subordinate group provided no information than when the subordinate group provided information encouraging high performance ratings. This

suggests raters did not perceive that the accuracy of their ratings will affect subordinate group approval of the performance ratings they provide for the bank teller. The valence associated with subordinate group approval was also not significantly different across the subordinate group manipulations. This implies that raters' motivational force to provide accurate performance ratings when the subordinate group provided information supporting high performance ratings was not driven by subordinate group approval of the ratings. Follow-up analyses, however, demonstrate that participants wanted to provide high performance ratings for the bank teller to a greater extent when the subordinate group information supported high performance ratings than when the subordinate group did not provide any information.

Performance ratings were not deflated, however, when the subordinate group provided information encouraging low performance ratings. The follow-up analyses provide some evidence that this may have been the case because the motivational force to provide deflated performance ratings was not stronger than the motivational force to provide accurate ratings. When the subordinate group provided information encouraging low performance ratings, raters' $P \rightarrow O$ expectancies that accurate ratings will result in subordinate group approval of ratings were not lower than when the subordinate group provided no information. Subordinate group approval was also not valued differently when the subordinate group provided information encouraging low performance ratings. Interestingly, participants in the subordinate group deflation information conditions indicated that they wanted to provide low performance ratings to a greater extent than participants in the no subordinate group information conditions.

Summary of Supervisor and Subordinate Group Interaction Effects

Consistent with expectancy theory and social impact theory, the findings suggest raters provide higher performance ratings when both the supervisor and subordinate group provide information encouraging high performance ratings than when no information is provided. Performance ratings were also higher when both the supervisor and subordinate group convey information supporting rating inflation than rating deflation. The results also indicate raters inflate performance ratings to a greater extent when information encouraging high performance ratings is conveyed by both sources than when the information is inconsistent. Specifically, results show ratings are higher when both the supervisor and subordinate group provide inflation information than when (a) the supervisor provides information encouraging low performance ratings and the subordinate group provides information encouraging high performance ratings, or (b) the supervisor provides information encouraging high performance ratings and the subordinate group provides information encouraging low performance ratings¹¹. Performance ratings are also higher when only the group provides information encouraging high performance ratings than when the group provides information encouraging high performance ratings and the supervisor provides information encouraging low performance ratings.

Contrary to predictions, performance ratings were not significantly different when both sources provided inflation information than when only the supervisor or only the subordinate

¹¹ This hypothesis was supported by the results of the One-way ANOVA. Specifically, the results of the One-way ANOVA indicate that performance ratings are higher when both the supervisor and subordinate group encourage high performance ratings than when the supervisor encourages high performance ratings and the subordinate group provides information encouraging low performance ratings. Tukey's test, however, indicates that the means were not significantly different. While the mean difference (.32) did not meet the cut-off (.388), it was close. This suggests that Tukey's test is a slightly more conservative test of the hypothesis than the One-Way ANOVA.

group provided inflation information. Ratings were also not significantly different when both provided information encouraging low performance ratings for the teller than when (a) only the supervisor provided information encouraging low performance ratings, (b) only the subordinate group provided information encouraging low performance ratings, (c) no information was provided, (d) the supervisor provided information encouraging high performance ratings and the subordinate group provided information encouraging low performance ratings, and (e) the supervisor provided information encouraging low performance ratings and the subordinate group provided information encouraging high performance ratings. Performance ratings were also not significantly different when the supervisor provided information encouraging low performance ratings and the subordinate group provided information encouraging high performance ratings than when the supervisor provided information encouraging high performance ratings and the subordinate group provided information encouraging low performance ratings.

Performance ratings were also not significantly different when the supervisor encouraged rating inflation and the subordinate group provided no information than when (a) the supervisor encouraged high performance ratings and the subordinate group provided information encouraging low performance ratings and (b) the supervisor encouraged low performance ratings and the subordinate group provided information encouraging high performance ratings. Ratings were also not significantly different when the subordinate group provided deflation information and the supervisor provided inflation information than when the subordinate group provided deflation information and the supervisor provided no rating information. Finally, ratings were also not significantly different when no information was provided than when (a) the supervisor encouraged high performance ratings and the subordinate group encouraged low performance

ratings and (b) the supervisor encouraged low performance ratings and the subordinate group encouraged high performance ratings.

Summary of Effect Magnitudes

While ratings were not significantly different when no information was provided and when the supervisor encouraged high (low) performance ratings and the subordinate group encouraged low (high) performance ratings, the effect of supervisor inflation and subordinate group deflation information on performance ratings was stronger than the effect of supervisor deflation information and subordinate group inflation information. Additionally, subordinate group information encouraging high performance ratings had a stronger effect on performance ratings than subordinate group information encouraging low performance ratings. Supervisor inflation information did not have a stronger effect on performance ratings than supervisor deflation information. Overall, however, inflation information had a stronger effect on performance ratings than deflation information.

Theoretical and Practical Implications

Overall, the results provide some evidence that social factors in the performance appraisal context affect rater motivation and performance ratings. This supports the rater goal perspective, which proposes that rater goals and subsequent performance ratings are affected by the performance appraisal context. While the present study only focused on the rater's motivation to provide accurate, high, and low performance ratings, rather than more specific rater goals (e.g., maintain group harmony or identify ratee strengths), the findings nonetheless provide support for rater goal theory. Specifically, contextual factors (i.e., the rater's supervisor and subordinate group) had effects on rater motivation and subsequent performance ratings.

By examining the interactive effects of key social contextual factors (i.e., the rater's supervisor and subordinate group) on rater motivation and performance ratings, the present study provides some insight as to what happens when raters encounter simultaneous information from different sources. This also answers calls for investigations focusing on the complex interpersonal relationships in the performance appraisal context. Specifically, results indicate that information from the supervisor interacted with subordinate group information to influence performance ratings as well as raters' valences of supervisor approval of performance ratings and $P \rightarrow O$ expectancies pertaining to supervisor approval of performance ratings resulting from accurate performance ratings. As expected, performance ratings were higher when both sources provided inflation information than when inconsistent information was provided. Consistent with social impact theory, conflicting information attenuated the effect of information on performance ratings, such that ratings were lower when the subordinate group and supervisor provided inconsistent information than when both provided inflation information. The attenuating effect is also in line with what was discussed by Mero and colleagues (2007). In their follow-up analyses, they found that raters accountable to both the ratees and the experiment administrator provided less accurate ratings than raters who were only accountable to the administrator. This is because raters inflated performance ratings when they were accountable to ratees. The current findings extend Mero and colleagues' (2007) findings to the rater's subordinate group. These findings are also consistent with performance appraisal research (e.g., Kozlowski et al., 1998) that suggests raters take relevant constituents into consideration when providing performance ratings.

The pattern of results in the present study also illustrates the importance of considering the complexities inherent in performance appraisal and avoiding simplistic conceptualizations of the context. For example, while not significantly different, the average performance rating in the

supervisor deflation information condition was higher than the average rating in the control condition. Considering the supervisor in isolation may suggest that supervisor deflation information either has no effect or that the effect may possibly be in the opposite direction. When the complexities in performance appraisal contexts are taken into account, a different understanding emerges. Specifically, when the supervisor encourages deflated ratings, and the subordinate group provides information encouraging inflated performance ratings, ratings are almost identical to those in the control condition (i.e., $N = 3.72$ versus $N = 3.74$, respectively). At the same time, when the subordinate group is the only salient source of information and provides information encouraging inflated performance ratings, ratings are significantly higher in the subordinate group inflation information condition ($N = 4.13$) than in the control condition ($N = 3.74$). Taken together, the deflation information from the supervisor appears to affect performance ratings in these more complex performance appraisal settings. This suggests that taking the complexity of the performance appraisal context into account is imperative, and studies considering these complexities are necessary, but currently lacking. Ultimately, considering these complexities inherent in performance appraisal provides a more realistic picture and a more nuanced and accurate understanding of the factors affecting performance ratings.

From a practical standpoint, it may be inappropriate for researchers to provide practical recommendations for improving performance appraisal accuracy based on studies that only take the effects of one contextual factor into account. It may also not be beneficial to generate simple remedies for improving performance appraisal accuracy based on research using overly simplistic and narrow conceptualizations of the performance appraisal context. As illustrated with the supervisor deflation information in the present study, it is problematic to examine the

effects of one social contextual factor in isolation. It is therefore important for future research to consider the effects of interactions between social contextual factors, and more generally, contextual factors, on performance appraisals. Additionally, providing simple remedies, without taking the complexities into consideration may result in other unanticipated problems. For example, solutions such as fixing rating scales or formats will not work as the performance appraisal context is not addressed by fixing these issues. Problems associated with forced distributions will not be resolved by improving the psychometric properties of rating instruments. Similarly, problems associated with social contextual factors in performance ratings cannot be resolved by focusing on factors such as rating purpose.

The findings of this study also speak to the different effects of information encouraging high versus low performance ratings, and rating inflation versus deflation, respectively. The results indicate that performance ratings were higher when both sources provided inflation information than deflation information. The average performance rating for the consistent inflation information condition ($M = 4.20$) was higher than the true score estimate ($M = 4$) and significantly higher than the average rating in the deflation information condition ($M = 3.77$). The rating in the deflation information condition was not significantly different than the mean rating in the control condition ($M = 3.74$). Furthermore, the effects of inflation information on performance ratings were generally stronger than the effects of deflation information on ratings. These findings provide empirical support for Longenecker and colleagues (1987) qualitative study. Specifically, raters appear to be reluctant to deflate performance ratings even when all the information they encounter encourages rating deflation. Clearly, more research focusing on rating deflation, and when it is likely to happen, is needed to inform our understanding of when raters are sufficiently motivated to provide deflated performance ratings.

Longenecker and colleagues' (1987) qualitative findings suggest that this will occur when the rater wants to send a message to the ratee or ensure that there is a paper trail for future dismissal. There is also the possibility that the rater will provide lower performance ratings than are merited because he or she does not like the ratee. Taken together with the present findings, this suggests that the rater's motivation to provide deflated performance ratings is largely affected by the ratee, as information from the rater's supervisor and subordinate group does not appear to have a strong effect on rating deflation. At this point, this is speculative and future research is needed to make strong conclusions. It would be beneficial to understand when and how social contextual factors, other than those related to the ratee, are likely to motivate raters to provide deflated performance ratings for a specific ratee. Understanding when raters are likely to provide deflated performance ratings will provide a more holistic understanding of the performance appraisal process.

As previously noted, results show that performance ratings were lower when the supervisor encouraged low performance ratings and the subordinate group encouraged high performance ratings than when only the subordinate group provided information encouraging high performance ratings. This also provides support for the attenuating effect and is consistent with social impact theory, which suggests that the lack of consensus reduces compliance (Latané & Wolf, 1981). It is also consistent with predictions generated from social impact theory, which suggest that the supervisor has greater influence than the subordinate group, given the supervisor's greater strength (i.e., power). This suggests that rating may actually be more accurate when raters encounter conflicting sources of information. Additional research is needed, however, before conclusions regarding conflicting information on performance rating accuracy can be drawn.

The predictions regarding attenuating effects were informed by integrating social impact theory with expectancy theory and the performance appraisal context. When considering social contextual factors, social impact theory can effectively be combined with expectancy theory to inform valences and develop hypotheses. For example, in the present study, it was predicted that the supervisor approval would have a higher valence than subordinate group approval as the supervisor was higher in strength compared to the subordinate group. A simple slopes *t*-test indicates that participants, across conditions, valued supervisor approval ($M = 3.69, SD = .75$) more than subordinate group approval ($M = 3.18, SD = .83$). This provides some evidence for the utility of social impact theory in the performance appraisal context. Furthermore, social impact theory was integrated with expectancy theory to predict that supervisor inflation information and subordinate group deflation information would have a stronger effect on performance ratings than supervisor deflation information and subordinate group deflation information. The results support this hypothesis, indicating that it may be beneficial to integrate these two theoretical frameworks to generate predictions.

One interesting finding was that information from the rater's supervisor and subordinate group had different effects on different motivational components. For example, the valence of supervisor approval of performance ratings was lower when the supervisor provided information encouraging low performance ratings than information encouraging high performance ratings. It appears that the type of information the supervisor provides affects how much supervisor approval of performance ratings is valued. Research (Jaworski & Kohli, 1991) suggests that positive supervisor feedback is both informational and motivational and is positively related to salesperson performance. Negative supervisor feedback, however, is informative but not particularly motivational and only minimally improves performance. It is possible that the

negative information was simply not sufficiently motivating. Supervisor information also had different effects on $P \rightarrow O$ expectancies. Specifically, $P \rightarrow O$ expectancies that accurate ratings will result in supervisor approval were higher when the supervisor provided no information or information encouraging high performance ratings than when information encouraging low performance ratings was provided. These findings suggest that it is important to look at the components separately to have a nuanced understanding of the effects of social contextual influences on rater motivation and performance ratings.

Surprisingly, information from the supervisor did not have main effects on performance ratings. Participants in the supervisor inflation information conditions reported a desire to provide high performance ratings to a greater extent than participants in the no supervisor information conditions, however. At the same time, raters' $\sum(P_{\text{inflate}} \rightarrow O) (V)$ were not different when the supervisor provided no information than when the supervisor conveyed a preference for high performance ratings. Valence of supervisor approval also did not change when the supervisor provided no information compared to when the supervisor provided information encouraging high performance ratings. These expectancy components suggest that the supervisor's information should have no effect on performance ratings. Consistent with expectancy theory, but contrary to the self-reports of a desire to provide high performance ratings, performance ratings were not significantly different when only supervisor inflation information was provided than when no information was provided. These findings provide some evidence that measuring expectancy components may be more useful than measuring self-reported motivation to engage in a specific behavior. The components appear to be a better predictor and may help us understand how the context affects different motivational components and subsequent behavior.

As predicted, performance ratings were higher when the subordinate group provided information encouraging high performance ratings than when no information was provided. Follow-up analyses indicate participants exposed to subordinate group inflation information wanted to provide high performance ratings to a greater extent than participants receiving no information from the subordinate group. Interestingly, subordinate group information did not have an effect on raters' $P \rightarrow O$ expectancies that accurate performance ratings result in subordinate group approval. Overall, these findings suggest that higher performance ratings in the subordinate group inflation condition may have been more a result of the kind of information that was conveyed by the subordinate group rather than a result of desiring subordinate group approval. Specifically, ratings may have been higher when the group provided inflation information as the information contained contextual performance information. The information conveyed by the subordinate group indicated that the bank teller was helpful and friendly and contributed positively to the work environment. Raters may consider this an important part of performance, despite the fact that this is not a part of task performance and the instructions were to provide task performance ratings. These findings are consistent with research demonstrating that contextual performance influences performance ratings and supervisory reward decisions (e.g., Dulebohn et al., 2005; Kiker & Motowidlo, 1999).

Taken together, these findings suggest that the subordinate group's information operated through an informational social influence route rather than a normative route. Specifically, as participants did not perceive that the accuracy of their performance ratings would affect subordinate group approval, normative social influence (i.e., conforming for approval) does not appear to be a factor. Rather, performance ratings were higher in the subordinate group inflation information condition than the no information condition due to informational social influence.

Informational social influence leads to conformity because information is accepted as evidence regarding reality (Deutsch & Gerard, 1954). It appears that raters provided higher performance ratings because they deemed the information as useful and relevant to the performance ratings. From a practical standpoint, organizations may have to consider what exactly is being evaluated when performance ratings are provided. The findings of this study suggest that any information raters deem relevant to performance, even when it is not task performance, are taken into consideration. The effect is stronger for positive information, however.

The relatively stronger effect of group information also has implications for multi-source, or 360-degree performance appraisals. Specifically, 360-degree appraisals generally involve the subordinates, supervisors, and peers of an employee (Mount, Judge, Scullen, Sytsma, & Hezlett, 1998). As noted by Bozeman (1997), while 360-degree feedback has been popular, it has also been criticized for lack of agreement across different sources. Murphy and Cleveland (1995) have pointed out that disagreement among raters may be expected as different sources will observe different behaviors. The present findings indicate that information that is communicated to raters, not just observed by the raters themselves, also affects the performance ratings they render for a specific employee. This raises some issues regarding the desirability of agreement across rating sources in 360-degree performance appraisal. It may be that considerable agreement across different sources is indicative of similar rater goals pertaining to the ratee, and not necessarily an indicator of the validity of the performance ratings.

Alternatively, it is possible that agreement is a fair indicator of the validity of performance ratings. Specifically, the rater's subordinate group (i.e., the ratee's peers) observes more contextual behaviors. The supervisor, however, may be more sensitive to task performance and less likely than the group to observe contextual performance. Given the importance of

contextual performance to organizational functioning, perhaps it is appropriate that the supervisor is made aware of contextual information related to a ratee.

The findings are also interesting in light of research (e.g., Wang et al., 2010) that has demonstrated raters' group harmony goals affect ratings. This implies that subordinate group approval of performance ratings is valued. Given the experimental setting of the present study, and absence of on-going relationships between the rater and the subordinate group, findings have to be taken with caution. Wang et al. (2010) used peer raters in Study 1, where individuals had ongoing relationships and were likely more concerned about maintaining interpersonal relationships than participants in the present study. In organizational settings, the subordinate group's information should be more likely to affect rater's $P \rightarrow O$ expectancies. In any case, additional research that focuses on mediating paths between the subordinate group as a social contextual factor and performance ratings are needed before strong conclusions can be drawn. If an expectancy framework is used, it is important to determine all relevant outcomes raters may consider and measure them. This will provide a more informed understanding regarding why social contextual factors, such as the rater's subordinate group, affect performance ratings.

A large number of the proposed hypotheses were not supported in the present study. This has a number of implications. On one hand, the null findings may be indicative of limitations with respect to the theory or study design, which are discussed in the following section. On the other hand, it may be that some of effects of social contextual factors that have been found in previous studies wash out when the complexities inherent in performance appraisal are taken into consideration. Additional research is needed to make conclusions regarding the source of the issue.

Limitations and Future Directions

The present study provides a conservative test (i.e., controlled experimental setting without real-life consequences) of the effects of social contextual factors on rater motivation and performance ratings. The experimental design did have a number of advantages, however. One, it allowed for the use of performance true score estimates, against which ratings could be compared. Two, it permits making causal inferences. Despite the conservative nature of the test, results indicate that these social contextual influences (i.e., the rater's supervisor and subordinate group) have an effect on performance ratings. In organizations, raters will experience consequences of the ratings they provide (e.g., Longenecker et al., 1987; Kozlowski et al., 1998). Additionally, the relationships in actual organizations are not temporary. Therefore, these effects are likely to be considerably stronger in field settings. Future research would benefit from providing more realistic settings, where raters must cope with real consequences resulting from the performance ratings they provide. Therefore, it would also be useful to test these predictions in more realistic experimental settings as well as field settings, to the extent that it is possible.

Furthermore, the performance ratings in the present study were provided for research purposes. Performance appraisal research (e.g., Wong & Kwong, 2007) shows that rating purpose affects performance ratings. For example, ratings given for research purposes are generally more lenient than appraisals conducted for administrative purposes (Harris, Smith, & Champagne, 1995). Considering the research purpose and experimental setting, social contextual factors in performance appraisal (i.e., the rater's supervisor and subordinate group) should have stronger effects on performance ratings in organizational settings than in the present study. This is because performance appraisals conducted in organizational settings are likely to be completed for administrative purposes and ratings generally have real-life consequences. Therefore,

supervisory performance ratings should be more susceptible to social contextual influences in organizational settings than in the present study.

It is also important to note that the present study used a sample of undergraduate business students. Research (e.g., Hall & Nougaim, 1968) suggests that managers' needs for achievement and affiliation increase as they advance in an organization. This suggests that raters in organizations, especially those who occupy higher positions in the organizational hierarchy, may be more motivated by information from supervisors and subordinate groups than the sample used for the present study. Future research is needed to adequately address this issue.

The research conducted on rater training suggests that the rater needs to have the ability to provide accurate performance ratings. In the present study, the rating task may have been challenging. It is therefore possible that ability moderates the relationship between the rater's motivation to provide accurate performance ratings and performance rating accuracy. As a result, it may be worthwhile to determine if raters' $E_{\text{accurate}} \rightarrow P_{\text{accurate}}$ expectancies (i.e., the subjective probability estimates that trying to provide accurate performance ratings will result in accurate performance ratings) moderate the relationship between rater motivation and rating accuracy. An alternative way to operationalize the ability component is by considering ratings provided for the practice scenarios, which were rendered prior to the experimental manipulations.

Given the null findings, it may be informative and beneficial to look at the effects of social contextual factors on ratings of each performance dimension, rather than overall performance ratings. For example, in this study, some performance dimensions (e.g., money handling) are more objective than others (e.g., professional conduct). It may be possible that

social contextual factors had stronger effects on performance ratings on the professional conduct dimension than the money handling dimension, as the former is more subjective in nature.

Furthermore, the way the motivational components were affected in the laboratory setting may not extend to the workplace. The follow-up analyses indicated subordinate group information did not affect the extent to which participants perceived accurate ratings result in subordinate group approval. In organizational settings, where relationships are on-going, the subordinate group's information may have a different effect on $P \rightarrow O$ expectancies and performance ratings. If this is the case in organizations, the subordinate group's information would affect the rater through normative and informational social influence, and the effect on ratings is likely to be stronger. If studies are conducted in field settings, it is important to identify the most important rater outcomes and determine how the expectancy components are affected by each and how they combine to affect the rater's motivational force to provide accurate versus distorted performance ratings.

While raters in the present study were aware of the implications their performance ratings have for the ratee, there was no information provided by the ratee. Research (e.g., Judge & Ferris, 1993; Longenecker et al., 1987) provides evidence that ratee factors other than performance (e.g., affect toward the subordinate) influence performance ratings. Therefore, to determine how ratee information, other than ratee task performance, interacts with subordinate group and supervisor information, future research should consider all three sources simultaneously. Additionally, research (e.g., Kozlowski et al., 1998; Spence & Keeping, 2010) suggests that the way peers evaluate their employees' performance affects the performance ratings rendered by the rater. As such, it would also be beneficial to investigate how information from the rater's peers influences ratings, when encountered simultaneously with information

from other key sources (e.g., rater's supervisor and subordinate group). Social impact theory would suggest that the rater will provide ratings that are in line with the source that has the greatest impact. These types of investigations, while challenging, would provide a more complete picture of the social intricacies inherent in performance appraisal.

Furthermore, the present study considered the effects of the rater's supervisor and subordinate group on performance ratings when the sources encouraged rating inflation or rating deflation. It would also be informative to understand what happens when these, and other sources, encourage accurate performance ratings. For example, are performance ratings more accurate when the rater's supervisor encourages accurate performance ratings and his or her subordinate group and peers encourage inflated performance ratings than when the supervisor expresses no preference and the peers and subordinate group encourage rating inflation? Understanding these and other complexities is an important first step in improving the accuracy of performance ratings.

The type of information conveyed by others in the performance appraisal context may also be important. In the present study, different types of information were provided by the subordinate group and the supervisor. The subordinate group's information contained facts about the ratee's contextual performance while the information from the supervisor was more of a self-interest manipulation. The type of information provided by each source is therefore confounded with the source of the information in the present study. Previous research (e.g., Spence & Keeping, 2010) has found that ratings were higher when it was advantageous for the rater. This was not found in the present study. Furthermore, supervisor information was overt in directing rating behavior. The information presented by the group was not. The overtness of the information is also confounded with the source of the information in the present study. Research

suggests that indirect information can be more influential (Raven, 2001). Therefore, future research is needed to disentangle the source of the information from the content of the information and the directness/indirectness of the information that is provided. Ferris and Mitchell's (1987) typology, focusing on intentional versus unintentional and active versus passive social cues, could provide a starting point for examining the different types of social influences raters may encounter when appraising employee performance. The type of social influence may play a role in the extent to which social influences affect performance ratings.

It would also be worthwhile to examine how the different types of information (i.e., information encouraging inflation or deflation) and different outcomes (i.e., positive or negative) affect rating inflation compared to rating deflation. Specifically, does inflation information have a stronger effect on performance ratings when providing inflated performance ratings is tied to positive outcomes compared to being tied to negative outcomes if ratings are not inflated? Research (e.g., Podsakoff, Todor, & Skov, 1982) on rewards and punishments would suggest that rewards have a stronger effect on behavior than punishments. This may hold true for performance rating behavior, however, future research should examine these complexities to further inform our understanding of the performance appraisal process. Additionally, the source of information or influence may also add complexity to this. Specifically, the rater's supervisor is more powerful than the rater, while the rater's subordinate group is not. It would be interesting to determine how the types of information and outcomes associated with the ratings that are rendered interact with the source that provides the information and outcomes.

Another limitation of the present study is that the ratings were provided for someone who is an average performer. Wang et al. (2010) found that ratings were inflated to a different extent for ratees with different performance levels. This suggests that inflation and deflation of

performance ratings for a ratee will be affected by the ratee's performance level. Taken together, this raises a number of questions. (1) When social contextual factors encourage rating inflation, will ratings be more inflated for a low performing compared an average performing employee? This has implications for employees and organizations. For example, if the low performer's ratings are inflated to a greater extent than the average performer's ratings, both employees may end up with comparable performance ratings. Given the fact that performance appraisals are used as a basis for other important decisions, such as merit pay raises, both the average and low performer may be receiving the same pay raises. This would be problematic as it would weaken the link between performance and rewards. Additionally, it would be interesting to determine if a high performing employee's performance ratings will be deflated to a greater extent than those of an average performing employee when social contextual factors encourage rating deflation. Again, this will have implications for ratees, rater's and the organization as a whole, given the extent to which the performance appraisal system is interconnected with other systems in the organization.

Additionally, the present study did not focus on individual differences. Random assignment should have ensured that experimental groups are equivalent on all individual differences. Future studies, however, should consider individual differences of raters. A number of individual differences may make raters more or less susceptible to the social influences encountered in the present study and those that are likely to be encountered in performance appraisal contexts. For example, when dealing with the social context of performance appraisal, individual differences such as agreeableness may come into play as some raters may be more susceptible to social influences than others. Another individual difference, Machiavellianism (e.g., Dahling, Whitaker, & Levy, 2009) may also affect the extent to which individuals are

influenced by social contextual factors. Machiavellianism reflects the extent to which individuals attempt to attain status and control over others, for example. Machiavellian individuals should therefore be more susceptible to the information provided by the supervisor in the present study than non-Machiavellian individuals. At the same time, some individuals may be more concerned with accuracy than accomplishing personal goals. These individuals are more likely to be affected by the information from the subordinate group, as this information may be perceived as being relevant to group and organizational functioning.

Judge and Ferris' (1993) findings also suggest the rater's affect toward the ratee affects performance ratings. This affective component was not assessed in the present study given the study design. It would be interesting, however, to examine the types of affective responses raters have to these various pressures and how they mediate the relationship between the information and performance ratings. Specifically, future research may benefit from understanding how the information changes the rater's affect toward the ratee and how that may influence future performance ratings.

It has been suggested that organizations need to focus on changing the context of performance appraisal, so that raters are motivated to provide accurate performance ratings. In the present study, raters encountered information that encouraged either rating inflation or deflation. It would be useful to examine if the accuracy of performance ratings may be improved when key social contextual factors (e.g., the rater's supervisor and subordinate group) provide information encouraging accurate performance ratings. It would also be informative to investigate what happens when one or more sources encourage accuracy. Additionally, it would be worthwhile to examine how the source(s) encouraging accuracy interact with sources encouraging rating inflation or deflation to affect performance ratings. Such studies may provide

insights regarding performance appraisal conditions that encourage raters to provide accurate performance ratings.

Lastly, the findings in this study suggest that social impact theory needs to be extended to include the types of social influence the target encounters from the source, rather than just focus on the source. If approval of the source of influence is a goal for the target, then influence is likely to operate through a normative route (e.g., Williams & Williams, 1989). If the target of the influence attempt is more concerned with being accurate, for example, then informational social influence should have a stronger effect (e.g., Deutsch & Gerard, 1954). Both Latané and Nida (1980) and Jackson (1986) have noted that social impact theory should be integrated with theories explicating the processes underlying social influence.

Summary

The findings of the present study suggest that raters provide higher performance ratings when both the supervisor and group provide information encouraging high ratings. Raters are reluctant to provide deflated performance ratings, however, even when both sources of information encourage low performance ratings. The results indicate that, overall, inflation information has a stronger effect on performance ratings than deflation information. Furthermore, the effect of information on performance ratings is attenuated when conflicting information is encountered. This suggests that both sources of influence are taken into account when raters rate performance. Finally, it appears that the group had more influence on the performance ratings, in this setting, due to informational social influence, as the information may have been perceived to be relevant for performance ratings. Overall, the findings provide some

support for expectancy theory and social impact theory as well as the relationship between the social context of performance appraisal and performance ratings.

From a practical perspective, the findings provide evidence that the complexities inherent in performance appraisal contexts interact to influence performance ratings. Therefore, it is important for future research to avoid narrow conceptualizations of the performance appraisal context and consider the inherent complexities. Researchers should also be cautious of providing practical recommendations based on the effects of isolated contextual factors on performance ratings. These simple solutions are likely to result in other, and potentially unanticipated issues.

REFERENCES

- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.
- Argo, J. J., Dahl, D. W., & Manchanda, R. V. (2005). The influence of a mere social presence in a retail context. *Journal of Consumer Research*, 32(2), 207-212.
- Athey, T. R., & McIntyre, R. M. (1987). Effect of rater training on rater accuracy: Levels-of-processing theory and social facilitation theory perspectives. *Journal of Applied Psychology*, 72(4), 567-572. doi:10.1037//0021-9010.72.4.567
- Austin, J. T., & Villanova, P. (1992). The criterion problem: 1917-1992. *Journal of Applied Psychology*, 77(6), 836-874.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.
- Banks, C. G., & Murphy, K. R. (1985). Toward narrowing the research-practice gap in performance appraisal. *Personnel Psychology*, 38(2), 335-345.
- Bernardin, H., & Buckley, M. (1981). Strategies in rater training. *Academy Of Management Review*, 6(2), 205-212. doi:10.5465/AMR.1981.4287782
- Bernardin, H. J., LaShells, M. B., Smith, P. C., & Alvares, K. M. (1976). Behavioral expectation scales: Effects of developmental procedures and formats. *Journal of Applied Psychology*, 61, 75-79.
- Bernardin, H. J., Tyler, C. L., & Villanova, P. (2009). Rating level and accuracy as a function of rater personality. *International Journal of Selection and Assessment*, 17(3), 300-310.
- Bocchiaro, P., Zimbardo, P. G., & Van Lange, P. A. M. (2012). To defy or not to defy: An experimental study of the dynamics of disobedience and whistle-blowing. *Social Influence*, 7, 35-50.
- Bommer, W. H., Johnson, J. L., Rich, G. A., Podsakoff, P. M., & MacKenzie, S. (1995). On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology*, 48(3), 587-605.
- Boswell, W. R., & Boudreau, J. W. (2002). Separating the Developmental and Evaluative Performance Appraisal Uses. *Journal of Business and Psychology*, 16(3), 391-412.
- Bozeman, D. P. (1997). Interrater agreement in multi-source performance appraisal: A commentary. *Journal of Organizational Behavior*, 18, 313-316.
- Burke, M. J., & Dunlap, W. P. (2002). Estimating interrater agreement with the average deviation index: A user's guide. *Organizational Research Methods*, 5(2), 159-172.

- Burke, M. J., Finkelstein, L. M., & Dusig, M. S. (1999). On average deviation indices for estimating interrater agreement. *Organizational Research Methods*, 2(1), 49-68.
- Campbell, J. J., Dunnette, M. D., Lawler, E. E., & Weick, K. E. (1970). *Managerial behavior, performance, and effectiveness*. New York, NY US: McGraw-Hill.
- Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, 55(1), 591-621. doi:10.1146/annurev.psych.55.090902.142015
- Cleveland, J., & Murphy, K. (1992). Analyzing performance appraisal as goal-directed behavior. In G. Ferris, & K. Rowland (Eds.), *Research in personnel and human resources management* (pp. 121-185). Greenwich, CT: JAI Press.
- Cleveland, J. N., Murphy, K. R., & Williams, R. E. (1989). Multiple uses of performance appraisal: Prevalence and correlates. *Journal of Applied Psychology*, 74(1), 130-135.
- Crandall CS, Eshleman A, O'Brien L. (2002). Social norms and the expression and suppression of prejudice: the struggle for internalization. *Journal of Personality and Social Psychology*, 82:359-78
- Cronbach, L. J. (1955). Assessment of individual differences. *Annual Review of Psychology*, 7(1), 173-196.
- Dahling, J. J., Whitaker, B. G., & Levy, P. E. (2009). The development and validation of a new Machiavellianism scale. *Journal of Management*, 35(2), 219-257.
- Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 1908, 8, 377-383.
- DeCotiis, T., & Petit, A. (1978). The performance appraisal process: A Model and some testable propositions. *Academy Of Management Review*, 3(3), 635-646. doi:10.5465/AMR.1978.4305904
- Deutsch, M., & Gerard, H. B. (1954). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, 51, 529-636.
- Dulebohn, J. H., Shore, L. M., Kunze, M., & Dookeran, D. (2005). The differential impact of OCBs and influence tactics on leader reward behavior and performance ratings over time. *Organizational Analysis*, 13(1), 73-90.
- Dunford, B. B., & Devine, D. J. (1998). Employment at-will and employee discharge: A justice perspective on legal action following termination. *Personnel Psychology*, 51(4), 903-934.

- Dunlap, W. P., Burke, M. J., & Smith-Crowe, K. (2003). Accurate tests of statistical significance for rwg and average deviation interrater agreement indexes. *Journal of Applied Psychology*, 88(2), 356-362.
- Farr, J. L., & Levy, P. E. (2007). Performance appraisal. In L. L. Koppes (Ed.), *The science and practice of industrial-organizational psychology: The first hundred years*. Mahwah, NJ: Erlbaum.
- Ferris, G. R., Judge, T. A., Rowland, K. M., & Fitzgibbons, D. E. (1994). Subordinate influence and the performance evaluation process: Test of a model. *Organizational Behavior and Human Decision Processes*, 58, 101-101.
- Ferris, G. R., & Mitchell, T. R. (1987). The components of social influence and their importance for human resources research. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resources management* (pp. 103-128). Greenwich, CT: JAI Press.
- Ferris, G. R., Munyon, T. P., Basik, K., & Buckley, M. R. (2008). The performance evaluation context: Social, emotional, cognitive, political, and relationship components. *Human Resource Management Review*, 18, 146-163.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2), 117-140.
- Fowler, R. L. (1987). A general method for comparing effect magnitudes in ANOVA designs. *Educational and Psychological Measurement*, 47, 361-367.
- French, J. R. P., & Raven, B. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies in Social Power*, p. 150-167. Ann Arbor, MI: University of Michigan Press.
- Gong, Y., Law, K. S., Chang, S., & Xin, K. R. (2009). Human resources management and firm performance: The differential role of managerial affective and continuance commitment. *Journal of Applied Psychology*, 94(1), 263-275.
- Greguras, G. J., Robie, C., Schleicher, D. J., & Maynard III, G. O. F. F. (2003). A field study of the effects of rating purpose on the quality of multisource ratings. *Personnel Psychology*, 56(1), 1-21.
- Group. 2013. In *Merriam-Webster.com*. Retrieved May 17, 2013, from <http://www.merriam-webster.com/dictionary/group>
- Hall, D. T., & Nougaim, K. E. (1968). An examination of Maslow's need hierarchy in an organizational setting. *Organizational Behavior and Human Performance*, 3(1), 12-35.
- Harris, M. M. (1994). Rater motivation in the performance appraisal context: A theoretical framework. *Journal of Management*, 20(4), 737-756. doi:10.1016/0149-2063(94)90028-0

- Harris, M. M., Smith, D. E., & Champagne, D. (1995). A field study of performance appraisal purpose: Research-versus administrative-based ratings. *Personnel Psychology*, 48(1), 151-160.
- Harris, K. J., Zivnuska, S., Kacmar, K., & Shaw, J. D. (2007). The Impact of Political Skill on Impression Management Effectiveness. *Journal of Applied Psychology*, 92(1), 278-285.
- Heneman, R. L. (1986). The relationship between supervisory ratings and results-oriented measures of performance: A meta-analysis. *Personnel Psychology*, 39(4), 811-826.
- Heneman, R. L., & Werner, J. M. (2005). *Merit pay: Linking pay to performance in a changing world*. Greenwich, CT: JAI Press.
- Ilgen, D. R., Barnes-Farrell, J. L., & McKellin, D. B. (1993). Performance appraisal process research in the 1980s: what has it contributed to appraisals in use? *Organizational Behavior and Human Decision Processes*, 54(3), 321. doi:10.1006/obhd.1993.1015
- Jacobson, D., & Eran, M. (1980). Expectancy theory components and non-expectancy moderators as predictors of physicians' preference for retirement. *Journal of Occupational Psychology*, 53, 11-26.
- Jackson, J. M. (1986). In defense of social impact theory: Comment on Mullen. *Journal of Personality and Social Psychology*, 50(3) 511-513.
- Jackson, J. M., & Latané, B. (1981). All alone in front of all those people: Stage fright as a function of number and type of co-performers and audience. *Journal of Personality and Social Psychology*, 40(1), 73-85.
- Jaworski, B. J., & Kohli, A. K. (1991). Supervisory feedback: alternative types and their impact on salespeople's performance and satisfaction. *Journal of Marketing Research*, 190-201.
- Judge, T. A., & Ferris, G. R. (1993). Social context of performance evaluation decisions. *Academy Of Management Journal*, 36(1), 80-105. doi:10.2307/256513
- Kane, J. (1994). A model of volitional rating behavior. *Human Resource Management Review*, 4, 283-310.
- Kane, J. S., & Bernardin, H. (1982). Behavioral observation scales and the evaluation of performance appraisal effectiveness. *Personnel Psychology*, 35(3), 635-641. doi:10.1111/j.1744-6570.1982.tb02215.x
- Kane, J. S., & Lawler III, E. E. (1979). Performance appraisal effectiveness: Its assessment and determinants. *Research in Organizational Behavior*, 1, 425-478.

- Kaplan, M. F., & Miller, C. E. (1987). Group decision making and normative versus informational influence: Effects of type of issue and assigned decision rule. *Journal of Personality and Social Psychology*, 53(2), 306-313. doi:10.1037/0022-3514.53.2.306
- Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations* (2nd ed.). New York, NY: John Wiley & Sons.
- Kiker, D. S., & Motowidlo, S. J. (1999). Main and interaction effects of task and contextual performance on supervisory reward decisions. *Journal of Applied Psychology*, 84(4), 602.
- Kline, T. B., & Sulsky, L. M. (2009). Measurement and assessment issues in performance appraisal. *Canadian Psychology*, 50(3), 161-171. doi:10.1037/a0015668
- Kozlowski, S.W.J., Chao, G.T., & Morrison, R.F. (1998). Games raters play: Politics, strategies, and impression management in performance appraisal. In James W. Smither, (Ed.) *Performance Appraisal: State of the Art in Practice* (pp. 163-205). San Francisco, CA: Jossey-Bass.
- Lance, C. E., Butts, M. M., & Michaels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, 9, 202-220.
- Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin*, 87(1), 72-107. doi:10.1037/0033-2909.87.1.72
- Larson, J. R. (1986). Supervisors' performance feedback to subordinates: The impact of subordinate performance valence and outcome dependence. *Organizational Behavior and Human Decision Processes*, 37, 391-408.
- Latané, B. (1981). The psychology of social impact. *American Psychologist*, 36(4), 343-356. doi:10.1037/0003-066X.36.4.343
- Latané, B. & Nida, S. (1980). Social impact theory and group influence: A social engineering perspective. In P. B. Paulus (Ed), *Psychology of group influence* (pp. 3-34). Hillsdale, N.J.: Erlbaum.
- Latané, B., & Wolf, S. (1981). The social impact of majorities and minorities. *Psychological Review*, 88(5), 438-53.
- Latham, G. P., & Wexley, K. N. (1977). Behavioral observation scales for performance appraisal purposes. *Personnel Psychology*, 30(2), 255-268. doi:10.1111/j.1744-6570.1977.tb02092.x
- Lawler, E., III. (1971). *Pay and organizational effectiveness: A psychological view*. New York, NY: McGraw-Hill.

- Lawler, E. E., III (1973). *Motivation in work organizations*. Monterrey, CA: Brooks/Cole Publishing Company.
- Lawler, E., III. (1990). *Strategic pay: Aligning organizational strategies and pay systems*. San Francisco, CA: Jossey-Bass.
- Lawler III, E. E., & Porter, L. W. (1967). The Effect of Performance on Job Satisfaction. *Industrial Relations*, 7(1), 20-28.
- Levine, H. Z. (1986). Performance appraisal at work. *Personnel*, 63(6), 63-71.
- Levy, D. A., Collins, B. E., & Nail, P. R. (1998). A new model of interpersonal influence characteristics. *Journal of Social Behavior and Personality*, 13(4), 715-733.
- Levy, P. E., & Williams, J. R. (2004). The Social Context of Performance Appraisal: A Review and Framework for the Future. *Journal of Management*, 30(6), 881-905. doi:10.1016/j.jm.2004.06.005
- Lissitz, R. W., & Green, S. B. (1975). Effect of the number of scale points on reliability: A Monte Carlo approach. *Journal of Applied Psychology*, 60, 10-13.
- Loi, R., Lai, J. M., & Lam, L. W. (2012). Working under a committed boss: A test of the relationship between supervisors' and subordinates' affective commitment. *Leadership Quarterly*, 23(3), 466-475. doi:10.1016/j.leaqua.2011.12.001
- Longenecker, C. O., Sims, J. P., & Gioia, D. A., (1987). Behind the mask: The politics of employee appraisal. *Academy Of Management Executive*, 1(3), 183-193. doi:10.5465/AME.1987.4275731
- Magee, J. C., & Galinsky, A. D. (2008). Chapter 8: Social hierarchy: The self-reinforcing nature of power and status. *Academy Of Management Annals*, 2351-398. doi:10.1080/19416520802211628
- Martin, D. C., Bartol, K. M., & Kehoe, P. E. (2000). The legal ramifications of performance appraisal: The growing significance. *Public Personnel Management*, 29(3), 379-405.
- Mathieu, J. E., Tannenbaum, S. I., & Salas, E. (1992). Influences of individual and situational characteristics on measures of training effectiveness. *Academy of Management Journal*, 35(4), 828-847.
- McClelland, D. C. (1961). *The Achieving Society*. New York: Van Nostrand Reinhold.
- McIntyre, R. M., Smith, D. E., & Hassett, C. E. (1984). Accuracy of performance ratings as affected by rater training and perceived purpose of rating. *Journal of Applied Psychology*, 69(1), 147-156. doi:10.1037//0021-9010.69.1.147

- Mero, N. P., Guidice, R. M., Brownlee, A. L. (2007). Accountability in the performance appraisal context: The effect of audience and form of accounting on rater response behavior. *Journal of Management*, 33(2), 223-252.
- Miner, J. B. (1984). The validity and usefulness of theories in an emerging organizational science. *Academy of Management Review*, 9(2), 296-306.
- Mitchell, T. R. (1983). The effects of social, task, and situational factors on motivation, performance, and appraisal. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), *Performance measurement and theory* (pp. 39-59). Hillsdale, NJ: Earlbaum.
- Mitchell, T. R., & Mickel, A. E. (1999). The meaning of money: an individual-difference perspective. *Academy of Management Review*, 24(3), 568-578.
doi:10.5465/AMR.1999.2202138
- Mohrman, A. M. and Lawler, E. E. (1983). Motivation and performance-appraisal behavior. In F. Landy and S. Zedeck (Eds.), *Performance measurement and theory* (pp. 173-189). Hillsdale, NJ: Erlbaum.
- Mount, M. K., Judge, T. A., Scullen, S. E., Sytsma, M. R., Hezlett, S. A. (1998). Trait, rater, and level effects in 360-degree performance ratings. *Personnel Psychology*, 51, 557-576.
- Mullen, B. (1985). Strength and immediacy of sources: A meta-analytic evaluation of the forgotten elements of social impact theory. *Journal of Personality and Social Psychology*, 48(6), 1458-1466.
- Murphy, K. R. (2008). Explaining the weak relationship between job performance and ratings of job performance. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 1(2), 148-160. doi:10.1111/j.1754-9434.2008.00030.x
- Murphy, K. R., & Cleveland, J. N. (1991). *Performance appraisal: An organizational perspective*. Needham Heights, MA US: Allyn & Bacon.
- Murphy, K., & Cleveland, J. (1995). *Understanding performance appraisal: Social, organizational and goal-oriented perspectives*. Newbury Park, CA: Sage.
- Murphy, K. R., Cleveland, J. N., Kinney, T. B., Skattebo, A. L., Newman, D. A., & Sin, H. P. (2003). Unit climate, rater goals and performance ratings in an instructional setting. *Irish Journal of Management*, 24(1), 48-65.
- Murphy, K. R., Garcia, M., Kerkar, S., Martin, C., & Balzer, W. K. (1982). Relationship between observational accuracy and accuracy in evaluating performance. *Journal of Applied Psychology*, 67(3), 320-325.
- Napier, N. K., & Latham, G. P. (1986). Outcome expectancies of people who conduct performance appraisals. *Personnel Psychology*, 39, 827-837.

- Noonan, L. E., & Sulsky, L. M. (2001). Impact of frame-of-reference and behavioral observation training on alternative training effectiveness criteria in a Canadian military sample. *Human Performance*, 14(1), 3-26.
- Osterman, P. (2007). Comment on Le, Oh, Shaffer and Schmidt. *Academy of Management Perspectives*, 3, 16–18.
- Pfeffer, J., Salancik, G. A., & Leblebici, H. (1976). The effect of uncertainty on the use of social influence in organizational decision making. *Administrative Science Quarterly*, 21(2), 227-245. doi:10.2307/2392044
- Pinder, C. C. (1998). *Work motivation in organizational behavior*. Upper Saddle River, NJ: Prentice Hall.
- Podsakoff, P. M., Todor, W. M., & Skov, R. (1982). Effects of leader contingent and noncontingent reward and punishment behaviors on subordinate performance and satisfaction. *Academy of Management Journal*, 25(4), 810-821.
- Porter, L. W., & Lawler, E. E. 1968. *Managerial Attitudes and Performance*. Homewood, IL: Irwin.
- Randall, R., & Sharples, D. (2012). The impact of rater agreeableness and rating context on the evaluation of poor performance. *Journal of Occupational and Organizational Psychology*, 85(1), 42-59. doi:10.1348/2044-8325.002002
- Raven, B. H. (1965). Social influence and power. In: I. D. Steiner and M. Fishbein (Eds.). *Current studies in social psychology* (pp. 371-382). New York: Holt, Rinehart & Wilson.
- Raven, B. H. (1992). A power/interaction model of interpersonal influence: French and Raven thirty years later. *Journal of Social Behavior and Personality*.
- Raven, B. H. (2001). Power/interaction and interpersonal influence: Experimental investigations and case studies. In: A. Y. L Chai and J. A. Bargh (Eds.). *The use and abuse of power: Multiple perspectives on the causes of corruption* (pp. 217-240). Michigan: Taylor & Francis.
- Reno, R. R., Cialdini, R. B., & Kallgren, C. A. (1993). The transsituational influence of social norms. *Journal of Personality and Social Psychology*, 64(1), 104-112. doi:10.1037/0022-3514.64.1.104
- Sanchez, R. J., Truxillo, D. M., & Bauer, T. N. (2000). Development and examination of an expectancy-based measure of test-taking motivation. *Journal of Applied Psychology*, 85(5), 739.
- Sedikides, C., & Jackson, J. M. (1990). Social impact theory: A field test of source strength, source immediacy and number of targets. *Basic and Applied Social Psychology*, 11(3), 273-281.

- Sheridan, J. E., Slocum, J. W., & Min, B. (1975). Motivational determinants of job performance. *Journal of Applied Psychology*, 60(1), 119.
- Sobieszek, B. I., & Webster Jr, M. (1973). Conflicting sources of evaluations. *Sociometry*, 550-560.
- Spence, J. R., & Keeping, L. M. (2010). The impact of non-performance information on ratings of job performance: A policy-capturing approach. *Journal of Organizational Behavior*, 31(4), 587-608.
- Spence, J. R., & Keeping, L. (2011). Conscious rating distortion in performance appraisal: A review, commentary, and proposed framework for research. *Human Resource Management Review*, 21(2), 85-95. doi:10.1016/j.hrmr.2010.09.013
- Spence, J. R., & Keeping, L. M. (2013). The road to performance ratings is paved with intentions: A framework for understanding managers' intentions when rating employee performance. *Organizational Psychology Review*, 1.-24.
- Stajkovic, A. D., & Luthans, F. (1997). A meta-analysis of the effects of organizational behavior modification on task performance, 1975-95. *Academy of Management Journal*, 40(5), 1122-1149. doi:10.2307/256929
- Stamoulis, D. T., & Hauenstein, N. A. (1993). Rater training and rating accuracy: Training for dimensional accuracy versus training for ratee differentiation. *Journal of Applied Psychology*, 78(6), 994-1003.
- Sulsky, L. M., & Balzer, W. K. (1988). Meaning and measurement of performance rating accuracy: Some methodological and theoretical concerns. *Journal of Applied Psychology*, 73(3), 497-506.
- Sulsky, L. M., & Day, D. V. (1994). Effects of frame-of-reference training on rater accuracy under alternative time delays. *Journal of Applied Psychology*, 79(4), 535-543.
- Sussmann, M., & Vecchio, R. P. (1982). A social influence interpretation of worker motivation. *Academy of Management Review*, 7(2), 177-186. doi:10.2307/257295
- Turner, N., Chmiel, N., Sandy Hershcovis, M., & Walls, M. (2010). Life on the line: Job demands, perceived co-worker support for safety, and hazardous work events. *Journal of Occupational Health Psychology*, 15(4), 482.
- Van Eerde, W., & Thierry, H. (1996). Vroom's expectancy models and work-related criteria: A meta-analysis. *Journal of Applied Psychology*, 81(5), 575-586.
- Van Scotter, J. R., & Motowidlo, S. J. (1996). Interpersonal facilitation and job dedication as separate facets of contextual performance. *Journal of Applied Psychology*, 81(5), 525.

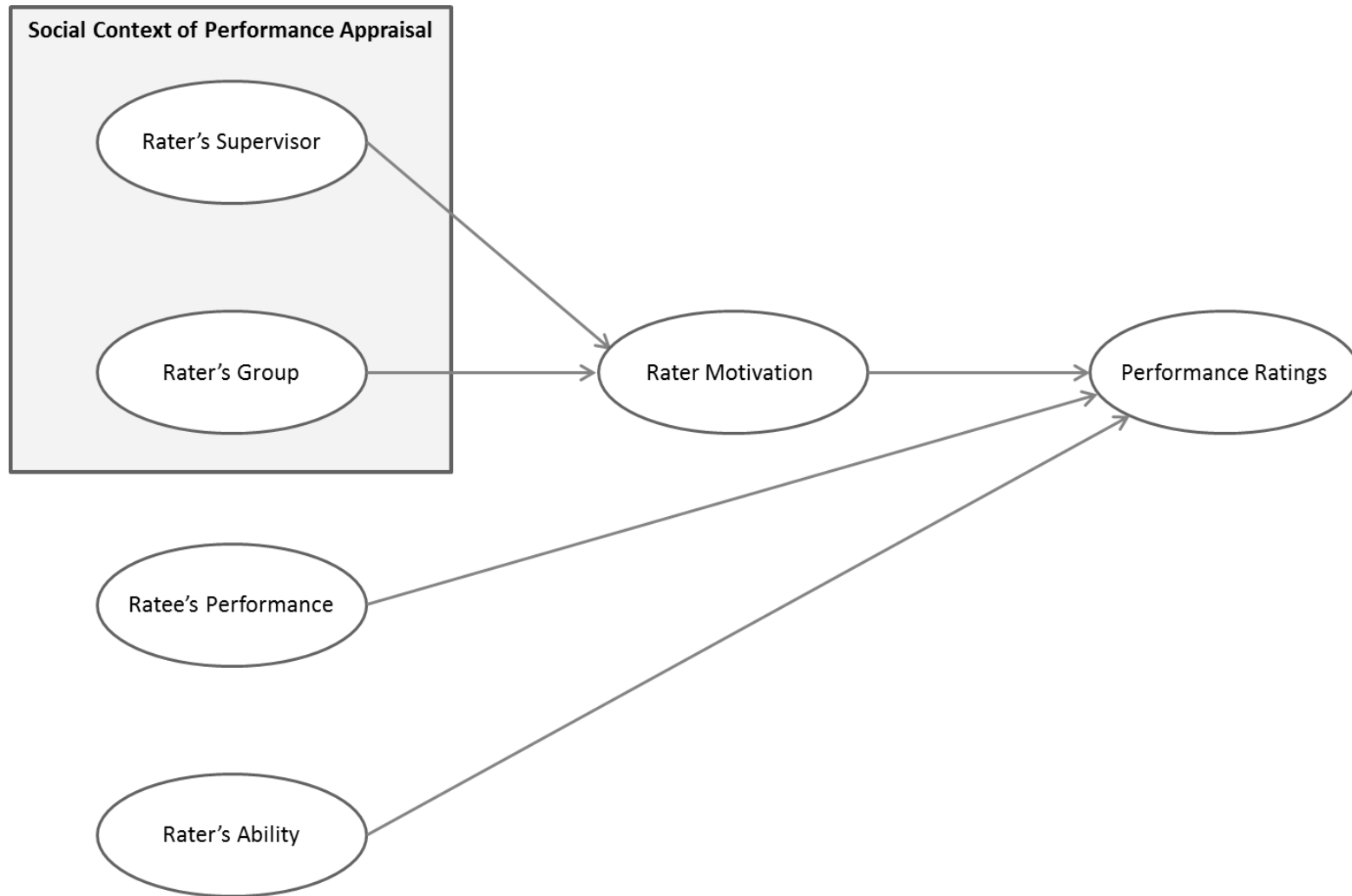
- Villanova, P., & Bernardin, H. J. (1989). Impression management in the context of performance appraisal. In R. A. Giacalone & P. Rosenfeld (Eds.), *Impression management in the organization* (pp. 299-314). Hillsdale, NJ: Erlbaum.
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81(5), 557-574.
doi:10.1037//0021-9010.81.5.557
- Vroom, V. H. (1964). *Work and motivation*. Oxford England: Wiley.
- Wang, X., Wong, K., & Kwong, J. Y. (2010). The roles of rater goals and ratee performance levels in the distortion of performance ratings. *Journal of Applied Psychology*, 95(3), 546-561. doi:10.1037/a0018866
- Wayne, S. J., & Ferris, G. R. (1990). Influence tactics, affect, and exchange quality in supervisor-subordinate interactions: A laboratory experiment and field study. *Journal of Applied Psychology*, 75(5), 487.
- Wayne, S. J., & Kacmar, K. M. (1991). The effects of impression management on the performance appraisal process. *Organizational Behavior and Human Decision Processes*, 48(1), 70-88.
- Werner, J. M., & Bolino, M. C. (1997). Explaining U.S. courts of appeals decisions involving performance appraisal: Accuracy, fairness, and validation. *Personnel Psychology*, 50(1), 1-24.
- Westaby, J. D., & Lowe, K. J. J. (2005). Risk-taking orientation and injury among youth workers: Examining the social influence of supervisors, coworkers, and parents. *Journal of Applied Psychology*, 90(5), 1027-1035. doi:10.1037/0021-9010.90.5.1027
- Wiesenfeld, B. M., Raghuram, S., & Garud, R. (2001). Organizational identification among virtual workers: the role of need for affiliation and perceived work-based social support. *Journal of Management*, 27(2), 213-229.
- Williams, K. D., & Williams, K. B. (1989). Impact of source strength on two compliance techniques. *Basic and Applied Social Psychology*, 10(2), 149-159.
- Woehr, D. J. (1994). Understanding frame-of-reference training: The impact of training on the recall of performance information. *Journal of Applied Psychology*, 79(4), 525.
- Woehr, D. J., & Huffcutt, A. I. (1994). Rater training for performance appraisal: A quantitative review. *Journal of Occupational And Organizational Psychology*, 67(3), 189-205.
doi:10.1111/j.2044-8325.1994.tb00562.x

Wong, K., & Kwong, J. Y. (2007). Effects of rater goals on rating patterns: Evidence from an experimental field study. *Journal of Applied Psychology*, 92(2), 577-585.
doi:10.1037/0021-9010.92.2.577

Zalesny, M. D. (1990). Rater confidence and social in performance appraisals. *Journal of Applied Psychology*, 75(3), 274-289.

APPENDIX A

Figure 1
Conceptual Model



APPENDIX B

Description of Participant's Role as Bank Manager

You are the manager of a local branch of a major American financial institution. Your duties include overseeing the day-to-day operations of the bank and serving as the main link between your branch and the corporate office. The job provides you with freedom to make many decisions, but also requires long hours at the office.

One of the most challenging aspects of your position is the need to ensure a smooth-running operation, while constantly struggling with limited resources. Your hard work has made the branch you manage one of the most successful ones, and that has not gone unnoticed by your supervisors. Specifically, your boss Pat (the district manager) attributes a large part of the branch's financial success to your outstanding managerial skills. In addition to impressing Pat, you have also succeeded in making the work environment positive for your subordinates. Your employees consider you a great boss and value all the things you do for them.

Recently, you have been under pressure from Pat to conduct the annual performance appraisals on the day-shift tellers at your branch. As with most of Pat's requests, you are expected to conduct these time-consuming evaluations without neglecting your other job responsibilities. Because your regular duties take up most of your work day, you don't have much extra time to conduct these assessments. You have decided that reviewing one teller each day of this week, and then submitting the completed evaluations next Monday, is the best approach to conducting the performance review process and completing it on time. Today, you have decided to review Julie.

Before you review Julie's performance, however, you will watch a training video. In the video, we will teach you what performance criteria you should base Julie's performance ratings on.

APPENDIX C

Scripts for Practice Performance Scenarios

Practice Scene 1:

Scene opens on a female TELLER and a male customer. The Teller is appropriately dressed in business casual attire, standing behind her station. Camera point of view (POV) shifts to the Teller. POV provides a clear view of the Teller from the waist up. On the tabletop, part of a digital financial calculator at the Teller's right side, part of a keyboard and mouse at the Teller's left side, along with the back of a flatscreen monitor (situated above the keyboard and mouse) and the a 1/4th rear view of the customer's face.

TELLER: Good morning! How may I help you?

CUSTOMER: I'd like to make a deposit.

TELLER: Sure thing, Mr. Anderson. I can help you with that.

CUSTOMER (*placing stack of bills on the countertop*): There it is.

TELLER (*taking bills, looks up at customer*): Do you have your deposit slip?

CUSTOMER (*removes deposit slip from shirt pocket*): Oh! Yes. Right here, sorry I forgot all about it.

Camera (POV) shifts to a top-down view of the teller's countertop. POV provides a clear view of the teller's hands and the materials placed there by the customer. Teller takes the bills from the countertop and begins to **quickly** count through them in a **clear, loud voice, separating** the money into different stacks according to their denomination with the bills **facing up**. Teller turns the deposit slip upright (so as to be read by the camera) and repeats the final tally. Having confirmed the total as correct, the teller collects all the bills and creates one stack.

Camera POV shifts to the dyad.

TELLER (*producing receipt from offstage and placing it on the countertop*): Thank you very much. Is there anything else I can help you with?

CUSTOMER: No, that's it. Thanks.

Scene ends.

Time: 24 seconds

Practice Scene 2:

Scene opens on a female TELLER and her female customer. The Teller is appropriately dressed in business casual attire, standing behind her station. Camera point of view shifts to the Teller. POV provides a clear view of the Teller from the waist up. At the Teller's right side, part of a keyboard and mouse at the Teller's left side, along with the back of a flatscreen monitor (situated above the keyboard and mouse) and the a 1/4th rear view of the customer's face.

TELLER: Is there anything else I can help you with, Ms. Baker?

CUSTOMER (takes receipt): No, Julie, that's everything.

TELLER: Before you go, I have a quick question for you. Are you aware your checking account balance is over \$8,000?

CUSTOMER (uneasy/hesitant): Yes . . . Uhm, I've been saving it for some time. Is something the matter?

TELLER (smiling/friendly): No, no, I'm sorry, I didn't mean to alarm you. But, you should consider opening a savings account instead of a checking account...

CUSTOMER (interrupts her, uncertain): Oh! I don't know. . . I like having access to my money whenever I need it.

TELLER (smiling): Okay! Just thought I'd mention it.

Scene ends.

Time: 35 seconds

Practice Scene 3:

Scene opens on a female TELLER and a male customer. The Teller is appropriately dressed in business casual attire, standing behind her station. Camera point of view shifts to the Teller. POV provides a clear view of the Teller from the waist up. On the tabletop, part of a digital financial calculator at the Teller's right side, part of a keyboard and mouse at the Teller's left side, along with the back of a flatscreen monitor (situated above the keyboard and mouse) and the a 1/4th rear view of the customer's face.

TELLER (*unenthusiastically*): Oh, hello.

CUSTOMER (*embarrassed/unsure*): Hello. I lost my debit card.

TELLER (*looks at him for a beat*): Mmm-hmm?

CUSTOMER: And . . . I'd like to get another one, please.

TELLER (*unenthusiastically*): Yeah, sure. Let me just get started on the paperwork . . . Weren't you in here last week?

CUSTOMER (*embarrassed, nervous*): Yes, I keep losing things like that. Kind of funny, right?

TELLER (*deadpan/serious for a beat, then rolls her eyes*): Hilarious.

CUSTOMER (*looking down*): Yeah, thank you. So . . . is there any paperwork I need to fill out, or anything?

TELLER (*continuing to type on the computer*): Just a second! I just need to cancel your *new*, new one. And yet again, order a . . . get you another one.

CUSTOMER (*whimpers*): Yeah . . . thank you. Sorry.

Scene ends.

Time: 47 seconds

APPENDIX D

Scripts for Performance Rating Scenarios

Rating Scene 1:

Scene opens on a female TELLER standing at her station, playing on her computer. She is appropriately dressed in business casual attire. A man enters the frame. Camera point of view is focused on the dyad, directed at the teller.

TELLER (*clicking her mouse*): Oh, hi! Sorry, I didn't hear you come in. Just a sec . . . Hang on, let me just finish this first . . . Hang on... Okay, there . . . Now, can I help you?

CUSTOMER (*places check and withdrawal slip on counter*): Hello. I'd like to make a deposit and a withdrawal please. [Customer sighs heavily] I have had the worst day today! I'll be so glad when it's over . . .

TELLER (*takes the check and turns to her computer*): Mmm-hmm?

CUSTOMER (*places check and withdrawal slip on counter*): Yeah, my next neighbor's dog kept me up, barking, all last night. . . And then, when I get into the office this morning, my secretary tells me that the client I've been romancing for two months has taken up with our competitor!

TELLER (*typing on her computer, disinterested*): Oh, really? I'm sorry to hear that.

Teller enters the information on the check into the computer. When finished, she picks up the withdrawal slip.

TELLER (*reading from the withdrawal slip*): So, Mr. . . Murray . . . would you like large or small bills?

CUSTOMER: Make it four twenties and the rest in fives, please.

Camera (POV) shifts to a top-down view of the teller's countertop. POV provides a clear view of the teller's hands and the materials placed there by the customer. Teller produces the bills from off-screen right and begins to **slowly and carefully** count through them in a **clear, loud voice, separating** the money into different stacks according to their denomination with the bills **facing up but not in the same direction**. The teller produces a receipt from off-screen right, turns it upright (for the camera) and repeats the final tally. Having confirmed the total as correct, the teller collects all the bills and creates one stack.

Camera POV shifts back to the dyad.

TELLER (*while pushing the cash across the countertop, coughs loudly*): Here you go, Mr. Murray. *COUGH COUGH*

CUSTOMER (*taken aback, collecting his money*): Jeez... thanks.

Customer finishes collecting his money and receipt and exits the frame.
Scene ends.

Time: 1 minute and 20 seconds

Rating Scene 2:

Scene opens on a female TELLER standing at her station having a personal conversation on her business phone. A male customer enters the frame. Camera point of view focuses on the dyad. The customer moves up to the window, but continues to talk on the phone.

CUSTOMER: Excuse me, miss?

TELLER (*putting her hand over the mouthpiece*): Yeah?

CUSTOMER: I, uhm . . . just want to make a withdrawal.

TELLER (*sighs heavily, still on the phone*): Uh, just a second. I've still got a couple minutes left on my break, but what the hell. [Speaking into the phone] Vanessa, I gotta help this guy, hang on a minute.

Teller presses a button on the business phone and hangs it up. She turns to face the customer, holding her hand out impatiently for his withdrawal slip.

TELLER: Give it here.

Customer hands her his withdrawal slip.

TELLER (*looking at the withdrawal slip, then up at the customer, eyebrows up*): Hey, buddy, are you sure you want to carry around this much cash? You want, like, an envelope or something?

CUSTOMER (*stammering*): Oh, no, it's fine. I'm going out of town and wanted to take some "just in case" money with me.

TELLER (*frowning*): Ok, well, no offense, but if you're planning on traveling, you don't want to have this much cash on your person. You'd be a walking lottery ticket for some lucky mugger. Seems kinda . . . irresponsible. Here's a better idea: Why don't I give you some traveler's checks, and then if something happens, you're covered, because the bank will replace those for a small fee. Something you can't do with cash. Sound good?

CUSTOMER (*thinking*): Actually . . . Yeah, that's a really good idea. Uh . . . thanks.

TELLER: No problem.

Teller turns to her computer and begins typing.

Scene ends.

Time: 1 minute and 09 seconds

Rating Scene 3:

Scene opens on a female TELLER standing at her station. She looks over as a female customer enters the frame. Camera point of view focuses on the dyad.

TELLER (*smiling*): Hello, miss. May I help you?

CUSTOMER (*placing a cashier's check and deposit slip on the countertop*): Yes. I'd like to make a deposit, please.

Teller takes the cashier's check and deposit slip off the countertop and turns toward her computer. She begins to enter the information into her computer, occasionally glancing at the check. After a few moments, she spots a discrepancy on the check and turns back to the customer.

TELLER (*frowning*): Ummmm, miss Perez, we have a problem. This check is expired. I cannot deposit this.

CUSTOMER (*shaking her head*): No, it is alright. It was a cashier's check I purchased a couple of months ago from my own account. It was supposed to be a graduation gift for my brother, but we all ended up chipping in and buying him a laptop instead.

TELLER (*uncertain*): Ok well. Again, I'm really sorry, but the check has expired. It looks like only 5 days ago. But once it has expired, we can't honor it.

CUSTOMER (*getting upset*): So what you're saying is that *my money* is, *expired*?! It was *my money*! I am just trying to deposit it back! What is the problem?

TELLER (*losing her patience*): Well, the problem, miss Perez, is that you're trying to deposit an expired check. Where it came from isn't important. Once it's expired, we cannot deposit it. I'm sorry, but that's final.

CUSTOMER (*reaches across and grabs the check out of Teller's hand*): This is ridiculous!! I'll—I'm gonna call your boss, and that's what I'm gonna do! Ridiculous!

Customer storms out. Teller watches her go, shrugs, and turns to her computer.

Scene ends.

Time: 1 minute and 11 seconds

Rating Scene 4:

Scene opens on a female TELLER and a customer. The Teller is appropriately dressed in business casual attire, standing behind her station. Camera point of view focuses on the dyad and then moves to the customer.

TELLER: Hello, Mr. Reynolds, how are you today?

CUSTOMER: Not so good, Julie. I need to make a change to one of my accounts. Can you help me with that?

TELLER: Sure. Depends on what it is that you need.

CUSTOMER: Unfortunately, Mrs. Reynolds and I are getting a divorce. Uhm, so I wanted to see if I can take her name off of our joint account. Is that something you help me with, Julie?

TELLER (startled): Gosh! I'm sorry! What happened?

CUSTOMER (embarrassed): I... I'd rather not talk about it.

TELLER: Oh ok. Well, geez, I'm sorry. That must be tough. Ummmm, well I don't know how to say, but I cannot do anything without Mrs. Reynolds present. So, if you want to take her name off, then you both need to be here, in person. I'm sorry, there's just no way around that.

CUSTOMER (frowning, puts check and deposit slip on counter): Oh... well let me just simply deposit this check and cash it.

TELLER (smiling): Ok I can help you with that!

Teller picks up the check and deposit slip and enters the information on the check into the computer.

TELLER (smiling): How would you like the \$500?

CUSTOMER (embarrassed): Mostly large bills. That'd be good.

Camera (POV) shifts to a top-down view of the teller's countertop. POV provides a clear view of the teller's hands and the materials placed there by the customer. Teller produces the bills from off-screen right and begins to **slowly** count through them in an **unclear, quiet voice, separating** the money into different stacks according to their denomination with the bills **facing up and in the same direction**. The teller produces a receipt from off-screen right, turns it upright (for the camera) and repeats the final tally. Having confirmed the total as correct, the teller collects all the bills and creates one stack.

Camera POV shifts back to the dyad.

TELLER (pushes the cash across the countertop): There you go, Mr. Reynolds. I'm sorry to hear about.... You know... sorry.

CUSTOMER (collects her money): Thanks, Julie.

Customer finishes collecting her money and receipt and exits the frame.

Scene ends.

Time: 1 minute and 23 seconds

Rating Scene 5:

Scene opens on a female TELLER standing at her station typing on a keyboard. A male customer enters the frame. Camera point of view focuses on the dyad. Teller greets the customer.

TELLER (*smiling*): Good afternoon, Mr. Julius, how can I help you today?

CUSTOMER (*frowning, holding an envelope and tossing it on the countertop*): First of all, you can refund these overdraft fees! Who charges this much for overdrafts? This is ridiculous! I've been banking here for years and this is how you guys treat me?

TELLER (*with friendliness and concern*): Uhm, let me look at the letter and see if there is anything I can do for one of our valued customers.

Customer hands her the letter and stands against the station, clearly irritated.

TELLER (*looking at letter, then entering information into the computer and looking at the screen*): I am really very sorry, Mr. Julius, but it looks like your account was overdrawn three times in the last month. Now we waived the non-sufficient fund fees twice already, which we do as a courtesy for our customers. But we are only allowed to do it twice. I am very sorry.

CUSTOMER (*angry*): Well that is unacceptable! I expect more from this bank!

TELLER (*patiently smiling and handing back the letter*): Again, I am very sorry t. I really wish I could do something. If you had overdraft protection, this would not have happened, but since you don't, there is nothing that I can do.

CUSTOMER (*shaking his head*): [Sighs heavily] Well, that's unacceptable! I, I, ... That's just not right!

TELLER: I understand, I really wish there was something I could do. If there's anything else I can help you with, don't hesitate to let me know.

CUSTOMER (*angry*): The only thing you could have helped me with, you didn't help me with, so no.

Customer just walks out without allowing the teller to respond.

Scene ends.

Time: 1 minute and 02 seconds

Rating Scene 6:

Scene opens on a female TELLER and her female customer. The Teller is appropriately dressed in business casual attire, standing behind her station. Camera point of view shifts to the Teller. POV provides a clear view of the Teller from the waist up. On the tabletop, part of a keyboard at the Teller's left side, along with the back of a flatscreen monitor and the a 1/4th rear view of the customer's face.

TELLER (*doodling on a piece of paper, bored*): Hey, there! Can I help you?

CUSTOMER (*nervous*): Hi, I'm uhm trying to uhm make a withdrawal, uhm, and also get some checks. I used all my checks to pay my monthly bills.

TELLER: Ok... Do you have your ID..., please?

CUSTOMER (*sorting through her bag and placing it on the counter*): Ugh, I can't seem to find it... I'm so sorry. (*Keeps digging through her purse*).

TELLER (*bored*): It's ok.

CUSTOMER (*finds ID and hands it to the teller*): Ah, here! I'm sorry about that.

Camera (POV) shifts to a top-down view of the teller's countertop. POV provides a clear view of the teller's hands and the materials placed there by the customer. Teller takes the bills from envelope and places them on the countertop and begins to **quickly** count through them in a **clear, loud voice, separating** the money into different stacks according to their denomination with the bills **facing up**. Teller turns the deposit slip upright (so as to be read by the camera) and repeats the final tally. Having confirmed the total as correct, the teller collects all the bills and creates one stack.

Camera POV shifts to the dyad.

TELLER (*producing receipt from offstage and placing it on the countertop*): Here you go, Ms. Bernard. I hope you have a better day.

CUSTOMER: Thanks so much.

TELLER (*smiling*): You're welcome.

Customer exits.

Time: 59 seconds

APPENDIX E

Performance Rating Scale

1. How would you rate Julie's performance on the following?

1 (Low) – 7 (High)

- a. ... the Money Handling dimension
- b. ... the Cross-Selling dimension
- c. ... the Professional Conduct dimension
- d. Julie's overall performance
- e. Counting money
- f. Explaining relevant bank services to the customer
- g. Having knowledge of and recommending the appropriate bank service to the customer
- h. Being a positive representative of the bank to the customer
- i. Dealing with cash
- j. Suggesting relevant bank services to the customer
- k. Acting in a professional manner

Money Handling: 1a, 1e, 1i

Cross-Selling: 1b, 1f, 1g, 1j

Professional Conduct: 1c, 1h, 1k

Overall Performance: 1d

APPENDIX F

Behaviorally-Anchored Rating Scale

Money Handling

7 (High Performance)

Julie counts the money out loud, quickly and accurately, making sure that every bill is placed in a neat pile by denomination, facing up and in the same direction.

6

5

Julie counts the money out loud and accurately, but she does it slowly. She also makes sure that every bill is placed in a neat pile by denomination, facing up but not in the same direction.

4 (Medium Performance)

Julie counts the money out loud, quickly and accurately, making sure that every bill is placed in a neat pile by denomination, facing up and in the same direction.

3

2

Julie counts the money accurately, but does so quietly and slowly, placing the bills into separate piles by denomination, with the bills facing up and down and in different directions.

1 (Low Performance)

Julie counts the money accurately, but does so quietly and slowly, placing all the bills into one messy pile.

Cross-Selling

Instructions: Please rate Julie's performance on the ***Cross-Selling*** dimension in the scene you just watched. Please use the rating rubric provided below. You may use your notes to rate as accurately as possible.

7 (High Performance)	Julie illustrates her knowledge about relevant bank services by suggesting and explaining the appropriate bank service to the customer.
6	
5	Julie illustrates her knowledge about relevant bank services by suggesting the appropriate bank service to the customer and how it could help, but fails to explain it fully.
4 (Medium Performance)	Julie illustrates her knowledge about relevant bank services by suggesting the appropriate bank service to the customer but failing to explain it.
3	
2	Julie illustrates her knowledge about relevant bank services by suggesting that if the customer had the appropriate bank service the situation would be different. However, she does not explain the service or even suggest it to the customer.
1 (Low Performance)	Julie illustrates a lack of knowledge about relevant bank services by failing to suggest the appropriate bank service to the customer.

Professional Conduct

Instructions: Please rate Julie's performance on the ***Professional Conduct*** dimension in the scene you just watched. Please use the rating rubric provided below. You may use your notes to rate as accurately as possible.

7 (High Performance)	Julie is business-like and friendly, uses appropriate language and makes the customer feel welcome.
6	Overall, Julie is business-like and friendly, and makes the customer feel welcome, but forgets to thank the customer at the end of the transaction.
5	
4 (Medium Performance)	Julie is friendly and uses appropriate language, but she is not being business-like by sharing too much information and making the customer feel somewhat uncomfortable rather than welcome.
3	While Julie uses appropriate language and is business-like, she is disinterested in what the customer is saying and appears unfriendly. She makes the customer feel unwelcome and doesn't even thank him for banking there.
2	
1 (Low Performance)	Julie is not friendly or business-like and uses inappropriate language. She is rude to the customer instead of making the customer feel welcome.

APPENDIX G

Graphic Rating Scale

1. Please rate Julie's performance on the following dimensions.
 - a. Counting the money out loud
 - b. Counting the money quickly
 - c. Placing bills in proper neat piles by denomination
 - d. Organizing the bills facing in the same direction
 - e. Making sure that all bills are facing up
 - f. Counting the money accurately
 - g. Being knowledgeable about relevant bank services
 - h. Suggesting appropriate bank services to customers
 - i. Explaining relevant bank services to customers
 - j. Being friendly
 - k. Being business-like
 - l. Making the customer feel welcome
 - m. Using appropriate language

2. Please indicate Julie's performance level on ...
 - a. ... the Money Handling dimension
 - b. ... the Cross-Selling dimension
 - c. ... the Professional Conduct dimension

3. Overall, Julie's performance for this scene is

All items were rated on a scale of 1 (Low) to 7 (High) with a "Not Applicable" option.

APPENDIX H

Formulas for Calculating Rating Distortion

Elevation

$$E^2 = (\bar{x}_{..} - \bar{t}_{..})^2$$

$$E = \sqrt{E^2}$$

$\bar{x}_{..}$ = mean rating across all rating scenarios and performance dimensions

$\bar{t}_{..}$ = mean true score across all rating scenarios and performance dimensions

Leniency

$$\text{Leniency/Severity}_k = \sum_{j=1}^r \frac{\frac{\sum_{i=1}^d (T_{ij} - R_{ijk})}{d}}{r}$$

d = number of dimensions

k = the k_{th} rater

r = number of rating scenarios

R = obtained rating

T = true score

APPENDIX I

Overview of Study for Participants

- I. Complete Questionnaire I
- II. Read about your role in this study
- III. Watch Training Video
- IV. Complete Questionnaire II
- V. Rate 3 Practice Scenes
- VI. Complete Questionnaire III
- VII. Read about the bank teller you will be rating
- VIII. Complete Questionnaire IV
- IX. Watch and rate 6 scenes of your teller
- X. Complete Questionnaire V
- XI. See Experimenter in Room 128

APPENDIX J

Experimental Manipulations

Supervisor Manipulation

No Rating Preference:

No information provided

Preference for Inflated Performance Ratings:

Earlier today, you had your daily phone conversation with your district manager Pat. During this conversation, Pat told you that you were being seriously considered for a long-overdue promotion. One thing that was holding up the promotion was the need to find a capable replacement for your current job. In the past, replacements have usually come from the day shift tellers. Pat thinks that Julie is the most promising candidate to replace you, but wants to wait to make the decision until after the current round of performance appraisals is completed. You got the impression that if you provide Julie with high performance ratings, then you would be promoted quickly and Julie named to replace you. If, on the other hand, you provide Julie with low ratings, then a thorough assessment of all the day-shift tellers would have to be conducted. This could delay your own promotion for quite some time. It would also mean a lot of extra work for you and Pat to identify a suitable alternative to fill your job.

Preference for Deflated Performance Ratings:

Earlier today, you had your daily phone conversation with your district manager Pat. In this conversation, you learned that there is an urgent need to reduce overhead in your branch. You both agree that the best way to cut costs is to lay off one of the day-shift tellers. Of the current staff, Julie has been with the bank for the shortest period of time. Pat feels that Julie is the best choice for the layoff, but wants to make the final decision after the current rounds of performance appraisals are completed. During the phone conversation, Pat emphasized that if you provide Julie with low ratings in her upcoming evaluation, the layoff process will be a simple matter. If you provide Julie with high ratings, however, the matter would be more complicated. It could take weeks to conduct a comprehensive review of all the day-shift tellers. You would be in charge of this time-consuming review, on top of having to manage the rest of your already-busy schedule.

Group Manipulation

No Rating Preference:

No information provided

Preference for Inflated Performance Ratings:

Members of Julie's work group, which consists of the other day-shift tellers, agree that she is universally well-liked, and friendly towards others. They note that she often goes out of her way to serve the difficult regular customers who most tellers try to avoid. She is always ready to lend a hand when her colleagues need assistance. The other tellers also note that they go to Julie for personal advice because she is a good listener and provides them with useful suggestions. They think Julie is very understanding and considers other people's points of view. She also comes up with great ideas for after-work gatherings and likes to attend them. They appreciate having her at the bank and feel that her presence contributes significantly to the positive work environment.

Preference for Deflated Performance Ratings:

Members of Julie's work group, which consists of the other day-shift bank tellers, agree they don't like her and that she is somewhat hostile to others. They note that she often goes out of her way to avoid dealing with difficult regular customers. She is usually reluctant to lend a hand when her colleagues need assistance. Her coworkers note that they do not go to Julie for any kind of advice because she does not treat personal conversations as confidential and shares private information with other co-workers. They think Julie is very opinionated and does not consider anyone else's point of view. She almost never contributes ideas for after-work gatherings and does not participate in them. They do not appreciate having her at the bank and feel that her presence has taken away from the positive work environment.

APPENDIX K

Formulas for Comparing Effect Magnitudes for ANOVA Designs

$$f_j^2 = [u_j F_j (v_j - 2) / v_j - u_j] / N_j$$

$$f_j^2 = \text{unbiased estimate of } \zeta^2$$

$$u_j = \text{degrees of freedom for numerator}$$

$$F_j = \text{observed variance ratio for the } j\text{th treatment}$$

$$v_j = \text{degrees of freedom for denominator}$$

$$N_j = u_j + v_j + 1$$

$$Z_j = [(1 - 2 / 9v_j) (u_j F_j / g_j)^{1/3} - (1 - 2h_j / 9g_j^2)] / [(2 / 9v_j) (u_j F_j / g_j)^{2/3} + (2h_j / 9g_j^2)]^{1/2}$$

$$Z_j = \text{approximation of } f^2$$

$$v_j = \text{degrees of freedom for denominator}$$

$$u_j = \text{degrees of freedom for numerator}$$

$$F_j = \text{observed variance ratio for the } j\text{th treatment}$$

$$h_j = u_j + 2N_j f^2$$

$$g_j = u_j + N_j f^2$$

$$N_j = u_j + v_j + 1$$

$$H_f = \sum_{j=1}^k Z_j^2$$

$$H_f = \text{approximate test for equality of } k \text{ effect sizes}$$

$$Z_j = \text{approximation of } f^2$$

APPENDIX L

Research Compliance Protocol Letters

November 14, 2008

MEMORANDUM

TO: Christopher Rosen
Nina Gupta
Trevor King
Emilija Djurdjevic

FROM: Ro Windwalker
IRB Coordinator

RE: New Protocol Approval

IRB Protocol #: 08-10-187

Protocol Title: *Pay and Performance Appraisal*

Review Type: ☒ EXEMPT ☐ EXPEDITED ☐ FULL IRB

Approved Project Period: Start Date: 11/14/2008 Expiration Date: 11/13/2009

Your protocol has been approved by the IRB. Protocols are approved for a maximum period of one year. If you wish to continue the project past the approved project period (see above), you must submit a request, using the form *Continuing Review for IRB Approved Projects*, prior to the expiration date. This form is available from the IRB Coordinator or on the Compliance website (<http://www.uark.edu/admin/rsspinfo/compliance/human-subjects/index.html>). As a courtesy, you will be sent a reminder two months in advance of that date. However, failure to receive a reminder does not negate your obligation to make the request in sufficient time for review and approval. Federal regulations prohibit retroactive approval of continuation. Failure to receive approval to continue the project prior to the expiration date will result in Termination of the protocol approval. The IRB Coordinator can give you guidance on submission times.

If you wish to make *any* modifications in the approved protocol, you must seek approval *prior to* implementing those changes. All modifications should be requested in writing (email is acceptable) and must provide sufficient detail to assess the impact of the change.

If you have questions or need any assistance from the IRB, please contact me at 120 Ozark Hall, 5-2208, or irb@uark.edu.

September 13, 2010

MEMORANDUM

TO: Christopher Rosen
Nina Gupta
Emilija Djurdjevic
Daanish Pestonjee

FROM: Ro Windwalker
IRB Coordinator

RE: New Protocol Approval

IRB Protocol #: 10-09-081

Protocol Title: *Generating Performance True Scores*

Review Type: ☒ EXEMPT ☐ EXPEDITED ☐ FULL IRB

Approved Project Period: Start Date: 09/13/2010 Expiration Date: 09/12/2011

Your protocol has been approved by the IRB. Protocols are approved for a maximum period of one year. If you wish to continue the project past the approved project period (see above), you must submit a request, using the form *Continuing Review for IRB Approved Projects*, prior to the expiration date. This form is available from the IRB Coordinator or on the Compliance website (<http://www.uark.edu/admin/rsspinfo/compliance/index.html>). As a courtesy, you will be sent a reminder two months in advance of that date. However, failure to receive a reminder does not negate your obligation to make the request in sufficient time for review and approval. Federal regulations prohibit retroactive approval of continuation. Failure to receive approval to continue the project prior to the expiration date will result in Termination of the protocol approval. The IRB Coordinator can give you guidance on submission times.

If you wish to make *any* modifications in the approved protocol, you must seek approval *prior to* implementing those changes. All modifications should be requested in writing (email is acceptable) and must provide sufficient detail to assess the impact of the change.

If you have questions or need any assistance from the IRB, please contact me at 120 Ozark Hall, 5-2208, or irb@uark.edu.

November 29, 2010

MEMORANDUM

TO: Christopher Rosen
Nina Gupta
Emilija Djurdjevic
Daanish Pestonjee

FROM: Ro Windwalker
IRB Coordinator

RE: New Protocol Approval

IRB Protocol #: 10-11-221

Protocol Title: *Performance and Performance Ratings*

Review Type: ☐ EXEMPT ☒ EXPEDITED ☐ FULL IRB

Approved Project Period: Start Date: 11/29/2010 Expiration Date: 11/28/2011

Your protocol has been approved by the IRB. Protocols are approved for a maximum period of one year. If you wish to continue the project past the approved project period (see above), you must submit a request, using the form *Continuing Review for IRB Approved Projects*, prior to the expiration date. This form is available from the IRB Coordinator or on the Compliance website (<http://www.uark.edu/admin/rsspinfo/compliance/index.html>). As a courtesy, you will be sent a reminder two months in advance of that date. However, failure to receive a reminder does not negate your obligation to make the request in sufficient time for review and approval. Federal regulations prohibit retroactive approval of continuation. Failure to receive approval to continue the project prior to the expiration date will result in Termination of the protocol approval. The IRB Coordinator can give you guidance on submission times.

If you wish to make *any* modifications in the approved protocol, you must seek approval *prior to* implementing those changes. All modifications should be requested in writing (email is acceptable) and must provide sufficient detail to assess the impact of the change.

If you have questions or need any assistance from the IRB, please contact me at 120 Ozark Hall, 5-2208, or irb@uark.edu.

November 7, 2011

MEMORANDUM

TO: Christopher Rosen
Nina Gupta
Emilija Djurdjevic
Daanish Pestonjee

FROM: Ro Windwalker
IRB Coordinator

RE: PROJECT CONTINUATION

IRB Protocol #: 10-11-221

Protocol Title: *Performance and Performance Ratings*

Review Type: ☐ EXEMPT ☒ EXPEDITED ☐ FULL IRB

Previous Approval Period: Start Date: 11/29/2010 Expiration Date: 11/28/2011

New Expiration Date: 11/28/2012

Your request to extend the referenced protocol has been approved by the IRB. If at the end of this period you wish to continue the project, you must submit a request using the form *Continuing Review for IRB Approved Projects*, prior to the expiration date. Failure to obtain approval for a continuation on or prior to this new expiration date will result in termination of the protocol and you will be required to submit a new protocol to the IRB before continuing the project. Data collected past the protocol expiration date may need to be eliminated from the dataset should you wish to publish. Only data collected under a currently approved protocol can be certified by the IRB for any purpose.

This protocol has been approved for 1,200 participants. If you wish to make *any* modifications in the approved protocol, including enrolling more than this number, you must seek approval *prior to* implementing those changes. All modifications should be requested in writing (email is acceptable) and must provide sufficient detail to assess the impact of the change.

If you have questions or need any assistance from the IRB, please contact me at 210 Administration Building, 5-2208, or irb@uark.edu.